



UNIVERSITY *of* LIMERICK

O L L S C O I L L U I M N I G H

Reader engagement in English, French and Spanish economics research articles: Contrasting questions in a comparable corpus of academic writing

By

Niall Curry

BA in Applied Languages, University of Limerick

MA in Teaching English to Speakers of Other Languages, University of Limerick

Student number: 0853224

A dissertation submitted to the School of Modern Languages and Applied Linguistics, in the Faculty of Arts, Humanities and Social Sciences at the University of Limerick, for the degree of Doctor of Philosophy

Supervisors: Dr Máiréad Moriarty & Prof. Emerita Angela Chambers

External Examiner: Prof. Shirley Carter-Thomas

Internal Examiner: Dr Anne O'Keeffe

Submitted to the University of Limerick on the 12th of November 2019

**Reader engagement in English, French and Spanish economics research articles:
Contrasting questions in a comparable corpus of academic writing**

Abstract

The aim of this thesis is to identify correspondences in questions used to engage readers in English, French, and Spanish economics research articles. Contrastive approaches to academic discourse occupy a small but growing space in the literature; however, the focus is typically on two languages. Moreover, most research in academic discourse has focussed on the English language, with relatively very little research available on questions as reader engagement devices in French and Spanish academic writing. This thesis asks to what extent the presence and functions of questions as reader engagement correspond in English, French, and Spanish economics research articles. Taking a corpus-based contrastive analysis approach, this thesis merges theories of contrastive linguistics and corpus linguistics to develop a framework of analysis that ensures effective comparability across academic languages. The corpus-based contrastive analysis approach is applied to the study of questions in the English and French economics subcorpora of KIAP (Fløttum *et al.* 2006), as well as a comparable Spanish subcorpus created for this study. Direct questions are identified through the use of a “?” while illocutionary force indicating devices are used to identify indirect questions. In the analysis, each direct and indirect question that serves to allow the writer to interact with the reader is analysed in terms of the following equivalences: frequency, function, type and form, location, passivity, tense and verbal modality, and question sentence type. A second analysis is presented in terms of these same equivalences; however the second analysis focuses on shared question function across languages. The contributions this thesis presents deliver new information on questions in economics research articles in each language. This thesis identifies how the aforementioned equivalences cause questions to differ across languages. The findings of this study contribute to linguistic descriptions of questions as reader engagement devices in economics research articles in each language, as well as research on contrastive studies of three languages. Moreover, the findings of this study allow for engagement with wider conversations on academic language in the multilingual academy, as well discussions surrounding the theoretical underpinning of corpus-based contrastive studies. In concluding this thesis, the findings are considered in terms of the applicability to the teaching and learning of academic language.

Declaration

I hereby declare that I am the sole author of this project. I certify that, to the best of my knowledge, my project does not infringe upon anyone's copyright nor violate any proprietary rights and that any ideas, techniques, quotations, or any other material from the work of other people included in my project are fully acknowledged in accordance with the standard referencing practices.

Signed: Niall Garry

Date: 12/11/2019

Acknowledgments

I would like to thank my external examiner, Prof. Shirley Carter-Thomas, my internal examiner, Dr Anne O’Keeffe, and my supervisors Dr Máiréad Moriarty and Prof. Emerita Angela Chambers for their support, mentorship, guidance, and time in helping me develop as a linguist, practitioner, and researcher. For use of data, I would like to thank Prof. Kjersti Fløttum, of the University of Bergen, for giving me access to the KIAP corpus. I would also like to thank the editorial board for the journals used to build my Spanish comparable subcorpus. These journals are: *Investigación Económica*, *La Revista de Economía Mundial*, and *El Trimestre Económico*. I would also like to thank the University of Limerick and Cambridge University Press for funding this PhD research.

Throughout this PhD, I have had the privilege to work among many dedicated communities of researchers and academics in different university contexts and all of these have played an important formative role in my development as a researcher. I would like to thank friends and colleagues at the University of Limerick, the University of Cambridge, Cambridge University Press, and Coventry University, where, at the time of submission, I work as a Lecturer in Academic Writing. My colleagues throughout these organisations are too many to name, but all have been so supportive and encouraging throughout this process. For this I am immeasurably grateful. I want to thank my friends who have accepted my absenteeism without judgement, who have read drafts of chapters, and who have listened to me talk through this work for the last three or so years. Your patience has been so valuable. In particular I want to thank Clare for helping me throughout and up to the very last submission, Pam and Seán for always supporting me, Robbie for being a sounding board and a reviewer, Geraldine for her insightful feedback, Claire for her mentorship, Delia for her editorial skills, Olivia for support throughout. I also want to thank the gang at 44, for listening to me talk about linguistics for the longest time. I want to thank my family for taking an interest in my work, my parents, and my sister Nyomi for helping me clean the data, and Mary, Christy, John, and Colette for taking an interest in my work, following my progress throughout this PhD, and supporting me and Chris in every way you could. Most of all, I want to thank Chris. I could not have done this without you and, as a linguist, I feel ironically lost for words to express how grateful I am to have you in my life. Thank you for supporting me, feeding me, and

making sure I mind myself and not just my work. I hope I can do the same in return for you, if you ever get inspired to do a PhD yourself.

Table of contents

Abstract	ii
Declaration	iii
Acknowledgments	iv
Table of figures	xvii
Table of tables	xix
Research outputs over the course of PhD research	xxv
Peer-reviewed papers, book chapters, and books	xxv
Selected talks, conference papers, and workshops on research related to PhD	xxv
1 Background and importance of research	1
1.1 Introduction	1
1.2 Background and research context	1
1.3 Importance and originality of this research	2
1.3.1 Academic discourse and academic writing in the multilingual context	3
1.3.2 Reader engagement in the economics research article in English, French, and Spanish	6
1.3.3 Corpus-based contrastive analysis	10
1.4 Research questions and delimitations	14
1.4.1 Research Questions	14
1.5 Delimitations	15
1.6 Chapter outline	16
2 Academic writing, reader engagement, and questions in English, French, and Spanish	21

2.1	Introduction	21
2.2	English academic discourse and the multilingual context	22
2.2.1	English as the lingua franca of academia: the impact on the multilingual academic context.....	24
2.2.2	Academic traditions and cultural identity: the impact on academic texts across languages.....	30
2.3	Disciplinary discourse communities	35
2.3.1	Focus disciplinary discourse communities.....	42
2.4	Genre in academic writing	43
2.4.1	The research article as a disciplinary genre	48
2.5	Academic writing and the English academic research article.....	51
2.5.1	Structure	51
2.5.2	Sentence length and sentence type	53
2.5.3	Passivity	55
2.5.4	Tense and verbal modality	56
2.6	Academic writing and the French academic research article.....	59
2.6.1	Structure	59
2.6.2	Sentence length and sentence type	60
2.6.3	Passivity, tense, and verbal modality	61
2.7	Academic writing and the Spanish academic research article	63
2.7.1	Structure, sentence length, and sentence type.....	63
2.7.2	Passivity, tense, and verbal modality	65
2.8	Metadiscourse in academic writing and the academic research article.....	66

2.9	Evaluation: writer stance and reader engagement in academic writing and the academic research article	72
2.9.1	Reader engagement in academic writing	75
2.10	Questions as reader engagement in academic writing and the academic research article	80
2.10.1	Question functions	80
2.10.2	Question types and question forms	86
2.11	Conclusion	92
3	Corpus-based contrastive analysis: background, theoretical foundation, and application to academic writing	95
3.1	Introduction	95
3.2	Contrastive linguistics: background and current state of the art	96
3.3	Theoretical foundations in contrastive linguistics.....	99
3.3.1	The <i>tertium comparationis</i> in contrastive linguistics.....	100
3.3.2	Theories of equivalence in contrastive linguistics	103
3.3.3	Challenges in measuring equivalence in contrastive linguistics: towards maximum similarity	108
3.4	Methodological frameworks in contrastive linguistics	110
3.5	Corpus linguistics: background, theoretical foundation, and relevance to contrastive linguistics.....	114
3.5.1	Corpus linguistics background and current state of the art	115
3.5.2	Corpora in corpus linguistics: theoretical underpinning and types of corpora	116
3.5.3	Methodological processes in corpus linguistics.....	119

3.5.4	Corpus linguistics and contrastive linguistics: multilingual corpora and complementary foundations	123
3.6	Corpus-based contrastive analysis, academic discourse, and the current state of the art	125
3.6.1	Corpus-based contrastive studies of academic discourse: a review	127
3.6.2	Gaps and future directions in corpus-based contrastive linguistics	133
3.7	Conclusion	134
4	A corpus-based contrastive analysis of direct and indirect questions in English, French, and Spanish	137
4.1	Introduction	137
4.2	Data: KIAP-EEFS	137
4.2.1	KIAP: comparable subcorpus of English and French economics research articles	138
4.2.2	Building a corpus of Spanish economics research articles comparable to KIAP	140
4.2.3	Finding questions in KIAP-EEFS	142
4.3	<i>Tertia comparationis</i> and equivalences: comparing and contrasting questions found in KIAP-EEFS	144
4.3.1	<i>Tertia comparationis</i>	144
4.3.2	Equivalence	144
4.3.2.1	Statistical equivalence: question frequency	145
4.3.2.2	Pragmatic equivalence: question function	145
4.3.2.3	Rule equivalence: question length	146
4.3.2.4	System equivalence: question type and form	146

4.3.2.5	Translation equivalence: question location in the research article ..	147
4.3.2.6	System equivalence: passivity in questions	147
4.3.2.7	System equivalence: tense and use of verbal modality in questions	148
4.3.2.8	Rule equivalence: sentence type within questions.....	148
4.3.3	Summary	148
4.4	Corpus-based contrastive analysis: question use as a shared rhetorical feature of reader engagement in English, French, and Spanish	149
4.5	Corpus-based contrastive analysis: question functions as a shared rhetorical feature of reader engagement in English, French, and Spanish	150
4.6	Identifying and addressing limitations	151
4.6.1	Data limitations	152
4.6.1.1	KIAP-EEFS	152
4.6.1.2	Author identity in KIAP-EEFS.....	153
4.6.1.3	Comparability challenges in KIAP-EEFS	154
4.6.2	Methodological limitations	154
4.6.2.1	Limitations of contrastive analysis approaches	154
4.6.2.2	Translation competence and judgement variables in corpus-based contrastive analysis	155
4.7	Closing remarks	156
5	Corpus-based contrastive analysis of question use as a shared rhetorical feature of reader engagement in English, French, and Spanish.....	158
5.1	Introduction	158
5.2	<i>Tertium comparationis</i> : question presence in KIAP-EEFS	158

5.3	Question functions	161
5.4	Question sentence length.....	164
5.5	Question type and form.....	165
5.5.1	Indirect questions: illocutionary force indicating devices	167
5.5.2	IFID: nouns	168
5.5.3	IFID: adjectives	169
5.5.4	IFID: verbs	169
5.6	Location in the text.....	172
5.7	Passivity, tense and verbal modality, and sentence type.....	173
5.7.1	Passivity	173
5.7.2	Tense and verbal modality	175
5.7.3	Sentence type	177
5.7.4	Passivity, tense and verbal modality, and sentence type: summary.....	178
5.8	Manifest similarities of questions based on degree.....	178
5.9	Identifying correspondence in question use as reader engagement	180
5.10	Conclusion	180
6	Corpus-based contrastive analysis of question function as a shared rhetorical feature of reader engagement in English, French, and Spanish	183
6.1	Introduction	183
6.2	Getting attention and focussing the reader	184
6.2.1	Getting attention and focussing the reader: frequency.....	184
6.2.2	Getting attention and focussing the reader: sentence length.....	184

6.2.3	Getting attention and focussing the reader: question type and form.....	185
6.2.4	Getting attention and focussing the reader: passivity, tense and verbal modality, and sentence type	187
6.2.4.1	Passivity	187
6.2.4.2	Tense and verbal modality and sentence type	188
6.2.5	Getting attention and focussing the reader: conclusion	190
6.3	Framing the discourse	190
6.3.1	Framing the discourse: frequency	190
6.3.2	Framing the discourse: sentence length	191
6.3.3	Framing the discourse: question type and form	192
6.3.3.1	IFID for indirect questions that frame the discourse	196
6.3.3.2	IFID: nouns and adjectives	196
6.3.3.3	IFID: verbs in indirect questions	197
6.3.4	Framing the discourse: location	199
6.3.5	Framing the discourse: passivity, tense and verbal modality, and sentence type	200
6.3.5.1	Passivity	200
6.3.5.2	Tense and verbal modality	201
6.3.5.3	Sentence type	202
6.3.6	Framing the discourse: conclusion.....	203
6.4	Organising the text	204
6.4.1	Organising the text: frequency	204
6.4.2	Organising the text: sentence length	205

6.4.3	Organising the text: question type and form	206
6.4.3.1	IFID for indirect questions that organise the text	210
6.4.3.2	IFID: nouns and adjectives	210
6.4.3.3	IFID: verbs in indirect questions	211
6.4.4	Organising the text: location	212
6.4.5	Organising the text: passivity, tense and verbal modality, and sentence type	212
6.4.5.1	Passivity	212
6.4.5.2	Tense and verbal modality	213
6.4.5.3	Sentence type	214
6.4.6	Organising the text: conclusion.....	215
6.5	Creating a niche.....	216
6.5.1	Creating a niche: frequency	216
6.5.2	Creating a niche: sentence length.....	217
6.5.3	Creating a niche: question type and form	218
6.5.3.1	IFID for indirect questions that create a niche.....	221
6.5.4	Creating a niche: location.....	222
6.5.5	Creating a niche: passivity, tense, polarity and sentence type	223
6.5.5.1	Passivity and tense and verbal modality	223
6.5.5.2	Sentence type	224
6.5.6	Creating a niche: conclusion	225
6.6	Expressing an attitude and counter-claiming	226

6.6.1	Expressing an attitude and counter-claiming: frequency	226
6.6.2	Expressing an attitude and counter-claiming: sentence length	226
6.6.3	Expressing an attitude and counter-claiming: question type and form ...	228
6.6.3.1	IFID for questions that express an attitude and counter-claim	230
6.6.4	Expressing an attitude and counter-claiming: location	231
6.6.5	Expressing an attitude and counter-claiming: passivity, tense and verbal modality, and sentence type	232
6.6.6	Expressing an attitude and counter-claiming: conclusion.....	233
6.7	Setting up claims and protecting the writer.....	234
6.7.1	Setting up claims and protecting the writer: frequency	234
6.7.2	Setting up claims and protecting the writer: sentence length.....	234
6.7.3	Setting up claims and protecting the writer: question type and form	235
6.7.3.1	IFID for indirect questions that set up claims and protect the writer	238
6.7.3.2	IFID: nouns and adjectives in indirect questions.....	238
6.7.3.3	IFID: verbs in indirect questions	239
6.7.4	Setting up claims and protecting the writer: location.....	239
6.7.5	Setting up claims and protecting the writer: passivity, tense and verbal modality, and sentence type	240
6.7.6	Setting up claims and protecting the writer: conclusion	243
6.8	Asking real questions	244
6.8.1	Asking real questions: frequency	244
6.8.2	Asking real questions: sentence length	244

6.8.3	Asking real questions: question type and form	245
6.8.3.1	IFID for indirect questions that ask real questions	247
6.8.3.2	IFID: nouns and adjectives in indirect questions.....	247
6.8.3.3	IFID: verbs in indirect questions	248
6.8.4	Asking real questions: location	248
6.8.5	Asking real questions: passivity, tense and verbal modality, and sentence type	249
6.8.6	Asking real questions: conclusion.....	251
6.9	Conclusion	252
7	Questions as reader engagement in English, French and Spanish: discussion of findings.....	254
7.1	Introduction	254
7.2	Academic writing and reader engagement in English, French and Spanish .	254
7.2.1	Reader engagement and questions in English, French, and Spanish	254
7.2.2	Function of questions in academic writing in English, French, and Spanish	258
7.2.3	Question sentence length in academic writing in English, French, and Spanish	266
7.2.4	Question types and forms in academic writing in English, French, and Spanish	268
7.2.5	Location of questions in the research article in English, French, and Spanish	275
7.2.6	Passivity in academic writing in English, French, and Spanish.....	278

7.2.7	Tense and modality in academic writing in English, French, and Spanish ..	280
7.2.8	Question sentence type in academic writing in English, French, and Spanish	283
7.2.9	Contribution to research on questions in academic writing in English, French, Spanish and contrastive linguistics	286
7.3	Contribution to research on the multilingual academy and corpus-based contrastive linguistics	286
7.3.1	The multilingual academy	286
7.3.2	Corpus-based contrastive linguistics	289
7.4	Conclusion	292
8	Implications and future directions	294
8.1	Introduction	294
8.2	Contributions: answering the research questions	294
8.3	Pedagogical implications	298
8.4	Looking forward	304
	Reference List	308
	Appendix A: Articles in specon subcorpus	353

Table of figures

Figure 3-1 The process of a contrastive functional analysis	113
Figure 5-1 Dispersion of questions in KIAP-EEFS in WPM	159
Figure 5-2 95% confidence intervals for question use across languages in WPM	160
Figure 5-3 Distribution of questions per function in WPM	161
Figure 5-4 Percentage distribution question functions in engecon	162
Figure 5-5 Percentage distribution question functions in frecon	162
Figure 5-6 Percentage distribution question functions in specon	163
Figure 5-7 Direct and indirect questions in KIAP-EEFS in WPM	165
Figure 5-8 Question types in KIAP-EEFS in WPM	166
Figure 5-9 Passivity in questions in KIAP-EEFS in WPM.....	174
Figure 5-10 Tenses in KIAP-EEFS in WPM	176
Figure 6-1 Distribution of questions per function in KIAP-EEFS in WPM.....	183
Figure 6-2 Question types and forms in getting attention and focussing the reader in KIAP-EEFS in WPM	186
Figure 6-3 Sentence type in getting attention and focussing the reader in KIAP-EEFS in WPM	189
Figure 6-4 Question types and forms in framing the discourse in KIAP-EEFS in WPM	193
Figure 6-5 Sentence types in framing the discourse in KIAP-EEFS in WPM	202
Figure 6-6 Question types and forms in organising the text in WPM in KIAP-EEFS ..	206
Figure 6-7 Tense and verbal modality in organising the text in KIAP-EEFS in WPM	214

Figure 6-8 Sentence types for questions that organise the text in KIAP-EEFS in WPM	215
Figure 6-9 Question types and forms in creating a niche in WPM in KIAP-EEFS	219
Figure 6-10 Question types and forms in expressing attitude and counter-claiming in WPM in KIAP-EEFS	229
Figure 6-11 Question types and forms in setting up claims and protecting the writer in WPM in KIAP-EEFS	236
Figure 6-12 Tense and verbal modality in setting up claims and protecting the writer in WPM in KIAP-EEFS	241
Figure 6-13 Sentence types in setting up claims and protecting the writer in WPM in KIAP-EEFS	242
Figure 6-14 Question types and forms in real questions in WPM in KIAP-EEFS	245
Figure 6-15 Sentence types in real questions in WPM in KIAP-EEFS	251

Table of tables

Table 2-1 Verbs used to embed questions in Karttunen (1977, pp.4-5)	87
Table 3-1 Comparison of types of equivalence in contrastive linguistics and translation studies (Chesterman 1998, p.38)	105
Table 3-2 Current types of equivalence in terms of Krzeszowski's seven types.....	106
Table 5-1 Questions in KIAP-EEFS	158
Table 5-2 Question length in words in KIAP-EEFS	164
Table 5-3 Question length ranges in KIAP-EEFS in WPM and percentage of all questions	164
Table 5-4 Percentage direct and indirect questions in KIAP-EEFS	166
Table 5-5 IFIDs that generate indirect questions in KIAP-EEFS in WPM and percentage of all questions	168
Table 5-6 Nouns used to create indirect questions in KIAP-EEFS in WPM.....	168
Table 5-7 Verbs used to generate indirect questions in KIAP-EEFS in WPM.....	169
Table 5-8 Questions per location in KIAP-EEFS in WPM and percentage of all questions	173
Table 5-9 Percentage passivity in KIAP-EEFS	174
Table 5-10 Percentage tenses in KIAP-EEFS	175
Table 6-1 Questions to get attention and focus the reader in WPM and percentage of all questions in KIAP-EEFS	184
Table 6-2 Getting attention and focussing the reader question length statistics in KIAP-EEFS	184
Table 6-3 Getting attention and focussing the reader question length in WPM and percentage of all questions in KIAP-EEFS	185

Table 6-4 Percentage direct questions in getting attention and focussing the reader in KIAP-EEFS	185
Table 6-5 Percentage direct questions in getting attention and focussing the reader per type and form in KIAP-EEFS	186
Table 6-6 Passivity in getting attention and focussing the reader in WPM and percentage of all questions in KIAP-EEFS	188
Table 6-7 Percentage sentence type in getting attention and focussing the reader in KIAP-EEFS	188
Table 6-8 Questions that frame the discourse in WPM and percentage of all questions in KIAP-EEFS	190
Table 6-9 Framing the discourse question length statistics in KIAP-EEFS	191
Table 6-10 Framing the discourse question length in WPM and percentage of all questions in KIAP-EEFS	191
Table 6-11 Percentage direct and indirect questions in framing the discourse in KIAP-EEFS	192
Table 6-12 Percentage direct and indirect questions in framing the discourse per type and form in KIAP-EEFS	193
Table 6-13 IFIDs that generate questions that frame the discourse in WPM and percentage of all questions in KIAP-EEFS	196
Table 6-14 Verbs that generate questions that frame the discourse in WPM in KIAP-EEFS	197
Table 6-15 Framing the discourse question location in WPM and percentage of all questions in KIAP-EEFS	199
Table 6-16 Passivity in framing the discourse in WPM and percentage of all questions in KIAP-EEFS	200

Table 6-17 Tense and verbal modality in framing the discourse in WPM and percentage of all questions in KIAP-EEFS	201
Table 6-18 Percentage sentence type in framing the discourse in KIAP-EEFS	202
Table 6-19 Questions that organise the text in WPM and percentage of all questions in KIAP-EEFS.....	204
Table 6-20 Organising the text question length statistics in KIAP-EEFS	205
Table 6-21 Organising the text question length in WPM and percentage of all questions in KIAP-EEFS.....	205
Table 6-22 Percentage direct and indirect questions in organising the text in KIAP-EEFS	206
Table 6-23 Percentage direct and indirect questions in organise the text per type and form in KIAP-EEFS.....	207
Table 6-24 IFIDs that generate questions that organise the text in WPM and percentage of all questions in KIAP-EEFS	210
Table 6-25 Verbs that generate questions that organise the text in WPM in KIAP-EEFS	211
Table 6-26 Organising the text question location in WPM and percentage of all questions in KIAP-EEFS.....	212
Table 6-27 Passivity in organising the text in WPM and percentage of all questions in KIAP-EEFS.....	212
Table 6-28 Percentage tense and verbal modality in organising the text in KIAP-EEFS	213
Table 6-29 Percentage sentence type in organising the text in KIAP-EEFS	215
Table 6-30 Questions that create a niche in WPM and percentage of all questions in KIAP-EEFS	217
Table 6-31 Creating a niche question length statistics in KIAP-EEFS.....	217

Table 6-32 Creating a niche question length in WPM and percentage of all questions in KIAP-EEFS	217
Table 6-33 Percentage direct and indirect questions in creating a niche in KIAP-EEFS	218
Table 6-34 Percentage direct and indirect questions in creating a niche per type and form in KIAP-EEFS	220
Table 6-35 IFIDs that generate questions that create a niche in WPM and percentage of all questions in KIAP-EEFS	221
Table 6-36 Verbs that generate questions that create a niche in WPM in KIAP-EEFS	222
Table 6-37 Creating a niche question location in WPM and percentage of all questions in KIAP-EEFS	223
Table 6-38 Passivity in creating a niche in WPM and percentage of all questions in KIAP-EEFS	224
Table 6-39 Tense and verbal modality in creating a niche in WPM and percentage of all questions in KIAP-EEFS	224
Table 6-40 Sentence type in creating a niche in WPM and percentage of all questions in KIAP-EEFS	225
Table 6-41 Questions that express attitude and counter-claim in WPM and percentage of all questions in KIAP-EEFS	226
Table 6-42 Expressing attitude and counter-claiming question length statistics in KIAP-EEFS	226
Table 6-43 Expressing attitude and counter-claiming question length in WPM and percentage of all questions in KIAP-EEFS	227
Table 6-44 Percentage direct and indirect questions in expressing and argument and counter-claiming in KIAP-EEFS	228

Table 6-45 Percentage direct and indirect questions in expressing attitude and counter-claiming per type and form in KIAP-EEFS	229
Table 6-46 IFIDs that generate questions that express attitude and counter-claim in WPM and percentage of all questions in KIAP-EEFS	230
Table 6-47 Verbs that generate questions that express attitude and counter-claim in WPM and percentage of all questions in KIAP-EEFS	230
Table 6-48 Expressing attitude and counter-claiming question location in WPM and percentage of all questions in KIAP-EEFS	231
Table 6-49 Passivity, tense and verbal modality, and sentence type in expressing and argument and counter-claiming in WPM and percentage of all questions in KIAP-EEFS	232
Table 6-50 Questions that set up claims and protect the writer in WPM and percentage of all questions in KIAP-EEFS.....	234
Table 6-51 Setting up claims and protecting the writer question length statistics in KIAP-EEFS	234
Table 6-52 Setting up claims and protecting the writer question length in WPM and percentage of all questions in KIAP-EEFS	235
Table 6-53 Percentage direct and indirect questions in setting up claims and protecting the writer in KIAP-EEFS	235
Table 6-54 Percentage direct and indirect questions setting up claims and protecting the writer per type and form in KIAP-EEFS.....	237
Table 6-55 IFIDs that generate questions that set up claims and protect the writer in WPM and percentage of all questions in KIAP-EEFS	238
Table 6-56 Setting up claims and protecting the writer question location in WPM and percentage of all questions in KIAP-EEFS	240
Table 6-57 Passivity in setting up claims and protecting the writer in WPM and percentage of all questions in KIAP-EEFS	240

Table 6-58 Percentage tense and verbal modality in setting up claims and protecting the writer in KIAP-EEFS	241
Table 6-59 Percentage sentence type in setting up claims in KIAP-EEFS.....	242
Table 6-60 Questions that as real questions and point forward in WPM and percentage of all questions in KIAP-EEFS.....	244
Table 6-61 Real questions question length statistics in KIAP-EEFS	244
Table 6-62 Real questions question length in WPM and percentage of all questions in KIAP-EEFS	244
Table 6-63 Percentage direct and indirect questions in real questions in KIAP-EEFS	245
Table 6-64 Percentage direct and indirect questions in real questions per type and form in KIAP-EEFS	246
Table 6-65 IFIDs that generate questions that ask real questions in WPM and percentage of all questions in KIAP-EEFS	247
Table 6-66 Nouns and adjectives that generate questions that ask real questions in WPM in KIAP-EEFS.....	248
Table 6-67 Real question location in WPM and percentage of all questions in KIAP-EEFS	249
Table 6-68 Passivity and tense and verbal modality in real questions in WPM and percentage of all questions in KIAP-EEFS	249
Table 6-69 Percentage sentence type in real questions in KIAP-EEFS.....	250
Table 7-1 Ranking of question functions in KIAP-EEFS	258

Research outputs over the course of PhD research

Over the course of my PhD I have had the opportunity to publish and present on my PhD research, as well as research in connected fields of corpus and language study. Each output offered opportunity for peer-review and valuable feedback and I am grateful to all who attended and reviewed my work. This section presents selected publications and those forthcoming, as well as talks that reflect my recent output.

Peer-reviewed papers, book chapters, and books

Curry, N. (2021 under-contract) *Academic writing and reader engagement: contrasting questions in English, French and Spanish corpora*, London: Routledge.

Curry, N. and Clark, T. (forthcoming) ‘Corpus-informed evaluation of examination criteria for MENA’, in McCallum, L. and Coombe, C., eds., *The assessment of L2 written English across the MENA region: a synthesis of practice*, Palgrave Macmillan.

Curry, N. (2019) ‘Learners as models: the pedagogical value of near peer role models’ *Speak Out! Journal of the IATEFL Pronunciation Special Interest Group*, 60, 34-43.

Curry, N and Chambers, A. (2017) ‘Questions in English and French Research articles in linguistics: a corpus-based contrastive analysis’ *Corpus Pragmatics*, 1(4), 327-350.

Selected talks, conference papers, and workshops on research related to PhD

Curry, N, (2019) Comparing academic writing in three languages: questions as metadiscourse in English, French, and Spanish, *International Conference on Corpus Analysis in Academic Discourse*, Jaume I University, Spain.

Curry, N. & Pérez-Paredes, P. (2019) From linguistic description to language learning: unlocking the value of the learner language in the Cambridge Learner Corpus, *International Corpus Linguistics Conference*, Cardiff University, UK.

Curry, N. (2019) How to write good academic and scientific English using corpora, *Escribir Ciencia en Inglés Universidad de Murcia*, University of Murcia, Spain.

- Curry, N., Dembry, C. & Pérez-Paredes, P. (2019) Workshop: English language corpora at Cambridge University Press, *Cambridge Language Sciences at the University of Cambridge*, University of Cambridge, UK.
- Curry, N. (2018) Rhetorical readers in academic writing: a corpus-based contrastive analysis of reader engagement in English, French and Spanish, *Cambridge Academic Online Conference*, University of Cambridge, UK.
- Curry, N. (2018) Rhetoric and reader pronouns in economics academic writing: a corpus-based contrastive analysis of English, French and Spanish research articles. *Inter-Varietal Applied Corpus Studies*, University of Valletta, Malta.
- Curry, N. (2017) Meeting the reader in academic writing: reader pronouns in English and French, *9th International Conference on Corpus Linguistics*, Paris 7, France.
- Curry, N. (2017) Corpus-based Contrastive analysis of reader engagement in academic writing: questions and reader pronouns, *International Research Methods Summer School*, University of Limerick, Ireland.
- Curry, N. (2016) A corpus-based contrastive analysis of question functions in English and French research articles, *Irish Association of Applied Linguistics*, Trinity College, Ireland.
- Curry, N. (2016) The form and function of questions in English and French research articles: a contrastive analysis, *Inter-Varietal Applied Corpus Studies*, Bath Spa University, UK.

Chapter One

Background and importance of research

1 Background and importance of research

1.1 Introduction

Research on academic writing comprises a dense canon within the field of academic discourse and applied linguistics. That being said, such research is dominated by a focus on the English language, with relatively very little research available on languages other than English. In fact, even within the study of academic writing in the English language, there remain a number of areas that are underdeveloped in the field. Relevant underdeveloped areas of research in academic writing pertain to studies of languages other than English, contrastive studies of more than two languages, disciplinary studies of economics research articles, and studies of reader engagement in academic writing. This thesis addresses these gaps through a corpus-based contrastive analysis that investigates the use of questions as devices of reader engagement in the economics research article in English, French, and Spanish. This chapter presents a detailed overview of this PhD thesis.

First, in Section 1.2, the background and context of the research is discussed briefly, outlining the contribution of this PhD to the academy. Next, in Section 1.3, the importance and originality of this research is outlined in greater detail, signalling the gaps in research that this thesis addresses. Following that, Section 1.4 presents the research questions that have guided this research and the delimitations that have focused it. Finally, a detailed outline of each chapter is presented in Section 1.5.

1.2 Background and research context

This PhD thesis is situated within the fields of academic discourse and corpus-based contrastive analysis. The research detailed in this thesis examines academic writing in English, French, and Spanish, with a focus on evaluative metadiscourse in the form questions as devices of reader engagement in the economics research article. Within this study of academic discourse, the focus herein is on the research article, which is arguably the ‘main channel of [...] scholarly communication’ (Holmes 1997, p.322). However, although it is clear from review of the existing literature that the research article has received considerable attention to-date, this research offers new perspectives of value to the current literature, owing to its focus on the discipline of economics and less-researched languages - French and Spanish. Moreover, the specific focus on evaluative metadiscourse

in the form of questions as devices of reader engagement and the detailed analysis of these questions in English, French, and Spanish addresses further gaps in the literature. This is because reader-oriented features of academic writing are less-studied in English, and even less so in French and Spanish.

The comparison and contrast of questions in academic research articles in English, French, and Spanish not only adds to the field of academic discourse, but it also adds to the field of contrastive linguistics. In contrastive linguistics, traditionally, most studies have focussed on comparing two languages, despite calls from the literature to incorporate three languages in contrastive studies (Van der Auwera 2012, p.75). In order to address these gaps in the literature on academic discourse and contrastive linguistics, a corpus-based methodology has been adopted, as its core tenets of representativeness and sampling reflect the importance placed on comparability and the *tertium comparationis* in contrastive linguistics. This approach takes form in a corpus-based contrastive analysis. The above-mentioned gaps and contributions are each considered in greater detail in the following section.

1.3 Importance and originality of this research

As mentioned in Section 1.2, this PhD is situated within the fields of academic discourse and corpus-based contrastive analysis; therefore, it contributes to both of these fields of scholarly research. To outline the importance and originality of this research to each field, Section 1.3.1 situates this research within its wider literary context and discusses research on academic discourse, its applications and uses, and the specific focus on academic writing in the research article. This is followed by a brief consideration of the wider global context, focussing on the role of English as a lingua franca of academia when compared to the role of other languages in academia. Next, in Section 1.3.2 the concept of discipline is discussed, identifying the importance of focussing on economics in this study. Following that, the research article, its language and interactional nature, and questions as reader engagement are discussed in order to pinpoint the gaps in the literature to which this thesis corresponds. Finally, Section 1.3.3 discusses the relationship between this thesis and the field of corpus-based contrastive linguistics.

1.3.1 Academic discourse and academic writing in the multilingual context

The field of academic discourse is well-established and broadly refers to the study of ‘the ways of thinking and using language that exist in the academy’ (Hyland 2011, p.171). Such a definition sees academic discourse emerge as a dynamic, flexible and complex discourse (Coffin and Donohue 2012, p.71; Heller and Morek 2015, p.179) with studies crossing genres, modes and languages, as well as text and context, metadiscourse and sentence-level studies of academic written and spoken language (Flowerdew and Wang 2015, p.82). In this study, the focus is on the written economics research article in English, French, and Spanish.

The research article is the most studied means of written academic communication, largely owing to its long-standing primacy for knowledge transfer in the academy (Holmes 1997, p.322). Its corollary, the conference paper, increasingly occupies an important space in the literature of spoken academic discourse (Rowley-Jolivet and Carter-Thomas 2005a; Rowley-Jolivet and Carter-Thomas 2005b), but, in comparison to written discourse, research on spoken academic discourse and research comparing written and spoken academic discourse is heavily underrepresented (Carter-Thomas and Jacques 2017, p.3). Further academic genres that have interested the community are those with which students engage, such as textbooks (Biber 2006a) or those which students produce, such as essays (Gardner and Nesi 2012). Studying these genres helps us to better understand the language presented to and used by novice members of the academic community. Overall, studying any of the above genres can yield important information on the language of the academy and its members.

While the research article has received more attention than these other genres, work on the research article is by no means complete. Given the focus of this thesis on three languages and the analysis of under-researched textual units of evaluation, namely questions, a firm theoretical foundation is required to avoid too many variables and unknowns in the research. Moreover, although the research article has received a lot of attention in studies of academic discourse, in the context of languages other than English and in the study of devices of reader engagement, there remains a need for further research. This is owing to the fact that research on academic discourse in languages other than English and reader engagement in general forms a disparate and sparse collection of

studies in the literature. While it is not possible to go into further detail, Section 2.2 and Section 2.10 respectively discuss in detail academic language in languages other than English and reader engagement in academic writing.

Overall, when considering academic writing within the field of academic discourse, research has followed two streams. The first refers to pedagogical student-centred analyses or linguistic descriptions to inform the education of learners of academic writing (Flowerdew 2002, p.1; Fløttum *et al.* 2006, p.26). The second is concerned with the description of academic writing to better understand the roles, identities, and practices of those in the academy (Hyland 2004a, p.1).

From a pedagogical perspective, much of the growth in research over the last 30 years has coincided with the development of English as the global lingua franca of academia (Dewey 2007; Seidlhofer 2010, p.356), which positions ‘non-Anglophone scholars [as] linguistically disadvantaged relative to native speaking academics when it comes to publication in English’ (Pérez-Llantada *et al.* 2011, p.45). Bocanegra-Valle argues that one potential reason for this disadvantage is the ‘insufficient command or inability to reach an acceptable academic style in that language’ (2014, p.1) by non-Anglophone scholars. This is supported elsewhere in the literature, as language policies of journals in economics, for example, have been found to impact journal identity, readership, and contributions (Henshall 2018). While this is a contentious matter, which has also been termed the ‘myth of linguistic injustice’ (Hyland 2016), the influence of first language on academic writing has largely been found to be, at the least, non-negligible (Carter-Thomas and Rowley-Jolivet 2013, p.112).

The linguistic disadvantage perspective has resulted in the development of a ‘multi-million dollar enterprise’ (Hyland and Hamp-Lyons 2002, p.1), the primary aim of which is to equip learners of academic English with rhetorical, interactive, and metadiscoursal competencies in academic writing through academic writing pedagogies (Hyland and Hamp-Lyons 2002, p.3). This industry has developed pedagogies in languages for specific purposes and English for academic purposes. These academic pedagogies extend to other languages such as *français langue académique* (Dervin *et al.* 2007, p.93) and *español con fines académicos* (Montemayor-Borsinger 2007). These are of important pedagogical value for the acculturation of students and the performance of academics within French and

Spanish university and academic contexts. However, academic English has become increasingly important in higher education outside of anglophone contexts (Klein and Boscolo 2016, p.311). This is largely owing to the role of English as a lingua franca of academia which has accelerated the need for the teaching and learning of English academic language (Dewey 2007, p.333; Seidlhofer 2010, p.366). Moreover, given the need for all to acculturate themselves within their respective academic communities, regardless of first language, there is clear use for research on academic discourse for informing the teaching of academic writing and academic literacies in fields like “writing across the curriculum” (WAC), “writing in the disciplines” (WID) and “academic literacies” (ACLIT) (McLeod and Miraglia 2001, p.10; Russell *et al.* 2009, p.410; Kruse 2013, p.39; Harper and Vered 2016, p.3; Tuck 2016, p.1612). However, while it is envisaged that the findings of this research will be useful for informing academic pedagogies¹, this research primarily adds to the second stream of research, which endeavours to better understand the linguistic practices of academic writers publishing in peer-reviewed journals in the discipline of economics in English, French, and Spanish.

The recognition of English as a global lingua franca has not only influenced global academic pedagogies. Rather, the growth of an educational industry that disseminates English language academic practices facilitates the hegemonic role of English as the lingua franca of academia in what is a multilingual academy. This has impacted the scientific field of academic discourse. It is clear that most studies on academic discourse have focused on the English language, due to its ‘privileged position [...] in the world of science today’ (Fløttum *et al.* 2006, p.18) with a much smaller canon of research centring on academic language in French and Spanish (Carter-Thomas and Jacques 2017, pp.2-3). This is unsurprising owing to the importance placed on the ability to use and engage with English-language academia and scholarship for the global construction of knowledge which has positioned knowledge of English academic discourse as a pinnacle of privilege (Hyland 2009, p.185).

The important contributions made through the medium of academic writing in languages other than English is well-documented in Ferguson (2007) and Bennet (2015), rendering academic discourse in French and Spanish a valuable discourse globally. However,

¹ The implications of the findings of this research to pedagogy are discussed in Chapter 8, Section 8.3.

Ferguson (2007) and Bennet (2015) also argue that owing to the growth of English as the lingua franca of academia, other academic languages are suffering from domain loss, diglossia, and potential epistemicides – concepts which are discussed in Section 2.2.1. Overall, language is an important variable in studies on academic discourse, with culture, in the form of academic traditions and cultural identity, well-documented as playing a significant role in informing the nature and linguistic features of academic writing in different languages. This pertains to concepts like reader- and writer-responsibility and content- and formal-oriented cultures – concepts which are discussed in Section 2.2.2.

The variation that occurs globally and culturally in academic writing in different languages raises important questions surrounding the comparability of academic discourse across languages. Therefore, there remains not only a need to add to the study of the academic discourse in languages other than English, in order to advance the field and to better understand the relationships that languages other than English have with the dominant English academic discourse. There is also a need to investigate the comparability of academic discourse in different languages, given the range of global and cultural variables that can render academic discourse in different languages incommensurable. The multilingual context of academic language is one area to which this thesis contributes, specifically in the context of English, French, and Spanish academic writing.

1.3.2 Reader engagement in the economics research article in English, French, and Spanish

The decision to focus specifically on the discipline of economics derives from trends that emerged in the mid-twentieth century, when research on academic writing began to consider discipline as an important variable (Prior 1998, p.196). This is a trend which has held strong today with extensive evidence of differences in linguistic practices across academic disciplines (Rongen Breivega *et al.* 2002, p.232; Salas 2015, p.20; Carter-Thomas and Jacques 2017, p.2). However, the concept of discipline is a loaded term, with political and cultural implications (Thompson 2001b, p.15) that make the process of defining what is contained within a discipline quite challenging. An example of this pertains to the fuzziness surrounding distinctions between a discipline and a sub-discipline (Darbellay 2015, p.170), which impacts the community captured within the

conceptualisation of discipline, and therefore the language captured within that discipline – see Section 2.3 for more detail on the subject of discipline.

Generally speaking, a discipline can be seen as a social community or ‘academic tribe’ (Becher 1989) that shares norms, bodies of knowledge, conventions, and modes of inquiry and culture (Hyland 2004a, p.8). Here the economics focus disciplinary discourse community is studied which is a discourse community that is not constrained by concepts such as time, language, and geographical location (Prior, 2003, p. 2). Instead, a focus discourse community is created on the basis of a shared focus of interest *i.e.* researching and disseminating knowledge in economics – see Section 2.3 for a more detailed discussion of discourse communities and Section 2.3.1 for a more detailed discussion of focus discourse communities. In the case of this research, the focus is on comparable research articles which are published in peer-reviewed journals in English, French, and Spanish in economics. As a discipline, economics has received relatively little attention in the literature when compared to the likes of linguistics or medicine. Therefore, while this study focuses on the economics disciplinary research article, it has been necessary to draw from research on similar and closely related disciplines to better understand that which is of most interest to this study: the role, organisation, and language of the research article.

The language of research articles is generally considered to be formal in nature and is quite often seen as objective or neutral (Hyland 2009, p.178). These two claims are now unpacked, beginning with formality. Academic writing in French and Spanish, like English, is generally considered to contain a specialised register that is formal in nature (Cassany 1999; Swales and Feak 1994; Salager-Meyer and Ariza 2004); however, the way in which that formality is expressed differs across languages (Biber and Conrad 2009, p.7). For example, the overall structure of the research article can differ across languages. Moreover, concepts of tense and passivity in academic contexts do not always share their typical use outside of the academy. For example, the past tense in English can serve to distance a researcher from a text with which they disagree, while the present serves to create a positive relationship between knowledge and information (Lakić 1997; Swales and Feak, 2012). Therefore, time serves as a proxy for implicit agreement or disagreement regardless of the temporal distance between the publications. Moreover, passivity is often

not related to hiding agency or a rhetorical decision but a necessity to achieve grammatical accuracy owing to the cotext (Flowerdew 2012). Other features of writing like modality and sentence length and complexity can differ across languages and disciplines. These features of academic writing are particularly important here as they have been used to analyse and compare questions across English, French and Spanish in order to better understand the use of questions in academic writing in each language. As such, they are discussed in detail in subsections within Section 2.5, 2.6, and 2.7, focussing on English, French, and Spanish respectively.

The second claim made of academic writing earlier refers to its neutrality and objectivity. However, this neutrality is increasingly questioned, with many researchers in the field focussing on the use of metadiscourse and, more specifically, the use of writer stance and reader engagement by academic writers to embed subjective and persuasive perspectives into their academic writing. Metadiscourse refers to the interpersonal features of discourse that position writers and readers within texts and help guide readers and organise texts – the interpersonal nature of writing and the use of metadiscourse forms the basis of the discussion of literature in Section 2.8. Metadiscourse does not contain propositional content per se but helps manage and organise that content in the construction of a narrative. Two particular evaluative features of interactional metadiscourse are stance and engagement. Stance refers to writers demonstrating their identity and often subjective evaluations in what are often seen as objective texts. Writer stance is often achieved through the use of hedges, boosters, self-mentions, and attitude markers (Hyland 2005a, p.177). Conversely, reader engagement refers to how writers include and position readers in texts in order to exploit their presence for the furtherance of the writers' narratives. This is often achieved through the use of directives, personal asides, shared knowledge, reader pronouns, and questions (Hyland 2005a, p.117) – stance and engagement receive greater attention in Section 2.9.

Much like studies of spoken language, where listenership is largely under-researched (Gardner 2001, p.1), the representation of readers in academic writing has been neglected in favour of the study of writer stance (Hyland 2005a, p.182; Gray and Biber 2012, p.18). While this dearth of studies on engagement is true for English, it is especially true in the context of economics academic writing and for languages other than English such as

French and Spanish. This is important since, against the backdrop of research on the cultural aspects of academic writing, French and Spanish writing is often positioned as less concerned with readers than English. This argument typically draws from research on reader-responsible or writer-responsible languages and content-oriented or formal-oriented languages across geographical and cultural boundaries. In terms of culture and academic writing, academic writing in romance cultures is considered quite distinct from English and less reader-oriented in general (Bennett and Muresan 2016, p.114); this concept is discussed in detail in Section 2.2.2. Overall, with the evident dearth of literature on metadiscourse and reader engagement in French and Spanish, it is imperative to undertake such research to help better understand the roles of writers and readers in academic writing across different academic cultures and languages, and contribute a further corpus-informed perspective to the established views of academic writing cultures. To address this need, this study has focussed on the use of questions as a device of reader engagement.

Questions as reader engagement in academic writing, which are discussed in detail in Section 2.10, have received little attention in the literature, with Hyland's work and framework for the categorisation of question functions arguably the most comprehensive in the context of academic writing (Orta 2010, p.81; McGrath and Kuteeva 2012, p.163; Rahimivand and Kuhi 2014, p.1495). Overall, questions are used in academic writing to directly engage with the reader through direct questions *i.e.* those signalled by a question mark, and to impart a question within the reader through indirect questions *i.e.* those signalled by an illocutionary force indicating device such as "we wonder" – for more on illocutionary force indicating devices for questions, see Section 2.10.2. Effectively, questions as reader engagement are direct and indirect questions that form part of the authorial text and that serve to ask the reader a question in order to exploit their position to manage and organise the discourse in some way. Questions can be used to 'hook readers' (Ruegg and Sugiyama 2013, p.13) and to re-open discussions (Soler 2007) for example, and Hyland's framework, which forms an important facet of this study, has identified seven functions of questions as reader engagement (Hyland 2002a) which are discussed in Section 2.10.1.

Overall, questions in academic writing have been studied in terms of their function (Thompson 2001a), their form (Blagojević and Misić-Ilić 2012) and their location in the research article (Ball 2009); however, such studies have not been undertaken at length. This thesis makes an important contribution to research on questions as reader engagement in academic writing by conducting and comparing two analyses of questions on three languages; these are presented in Chapter 5 and Chapter 6. Building on Hyland's framework (2002a) and enriching it with more recent studies from the field, the following seven functions are used to categorise question functions here:

- getting attention and focussing the reader;
- framing the discourse;
- organising the text;
- creating a niche;
- expressing an attitude and counter-claiming;
- setting up claims and protecting the writer;
- and asking real questions and pointing forward.

So far, the contributions of this thesis to academic metadiscourse, reader engagement and questions have been discussed. In the next section, the focus is on a further contribution that this research makes to the field of contrastive linguistics and corpus-based contrastive analysis.

1.3.3 Corpus-based contrastive analysis

As previously stated in Section 1.3.1, research on the rhetorical functions of reader engagement in academic writing has largely focused on English, with fewer studies focusing on languages other than English. Fewer still take a contrastive perspective, comparing the rhetorical features of research articles across languages. Contrastive linguistics allows for such studies and is generally underpinned by two core concepts: the *tertium comparationis* and equivalence, both of which are discussed in Section 3.3.

In brief, the *tertium comparationis* is the assumed common ground on which the contrastive analysis is based (Connor and Moreno 2005, p.156), while equivalence refers to the criterion or criteria used to test the *tertium comparationis* (Chesterman 1998, p.29). There can only be one *tertium comparationis* per analysis and as this thesis contains two

analyses there are two *tertia comparationis*, which are analysed according to several established equivalences which will be outlined later in this section. The goal of a contrastive analysis is to find out how the *tertium comparationis* corresponds across languages. Often, there are multiple options to choose from when determining corresponding language and expressions in a contrastive analysis, and these are called the manifest similarities; those manifest similarities that best fit the context are named correspondences. Contrastive studies can be further categorised as either convergent or divergent. Convergent analyses address a comparable feature in each language, while divergent analyses address a feature in one language and seek comparable features in other languages. This study presents a convergent analysis.

As will be discussed in Section 1.4, this thesis and its analyses are guided by a number of research questions. In the first analysis, which addresses the first research question and is presented in Chapter 5, the *tertium comparationis* is the assumption that questions used to engage the reader are present in each language. These questions are identified and analysed in terms of the following equivalences: frequency, function, length, type and form, location, passivity, tense and verbal modality, and sentence type. The second analysis, which addresses the second research question and is presented in Chapter 6, presents question function as the *tertium comparationis*; this categorises questions according to their perceived function. The questions in each function are then analysed in terms of the following equivalences: frequency, length, type and form, location, passivity, tense and verbal modality, and sentence type. These two analyses allow for a better view of questions in general in each language and for the capture of any idiosyncratic features of individual functions of questions in each language. Both analyses can be described as convergent analyses as questions are identified in each language and are then brought together to be compared. This, of course, limits the possibility to finding questions as being similar to questions and not to other devices of reader engagement such as directives or reader pronouns that a divergent analysis may facilitate. Both approaches are limited; however, a continuation of this study could be a comparison with divergent analyses of questions.

Returning to the *tertium comparationis*, it is important to note the flexibility of the concept as it not only refers to the linguistic item being compared, but it also captures the need to

ensure comparability at all levels of analysis. This means that there is a need to make assurances that the data being compared are comparable and accurately reflect samples of populations. As such, corpus linguistics as a methodology is an important aspect of this study which makes use of a corpus-based contrastive analysis approach.

Corpus linguistics is an approach in applied linguistic research that supports the empirical investigation of language (Biber and Reppen 2012, p.ix) using corpus analysis software and large collections of texts called corpora. These software allow for a range of methodological processes, discussed in Section 3.5.3, that allow for the identification and analysis of discrete parts of language and their accompanying metadata. The corpus-based contrastive analysis detailed in this thesis uses data from two corpora. The first is the KIAP² corpus (Fløttum *et al.* 2006), which is a multilingual comparable corpus composed of 450 research articles, with 150 each in English, French and Norwegian. These research articles are sub-categorised according to discipline with 50 in linguistics, economics, and medicine in each language. This thesis draws on the English and French economics subcorpora from the KIAP corpus. For the Spanish data, a comparable Spanish economics corpus was compiled for this study. This corpus is named specon³. Together these three subcorpora make up KIAP-EEFS which represents KIAP economics English, French, and Spanish. KIAP-EEFS can be described as a multilingual, comparable, ‘finite-size, machine-readable corpus’ (Fløttum *et al.* 2006, p.7).

The components of KIAP-EEFS, both in terms of language and discipline, were chosen for a number of reasons. The choice of language is informed by English being a known lingua franca (Seidlhofer 2010) and French and Spanish being major world languages. Previous research in contrastive linguistics has focused most often on English and another major language, such as Lorés-Sanz *et al.*’s (2010) work on English and Spanish, or English and a minor language such as Jafari’s (2014) work on English and Persian. For trilingual studies, research typically compares the lingua franca to a major and a minor language, as in Fløttum *et al.*’s (2006) work on English, French, and Norwegian or Ciapuscio and Otañi’s (2002) work on English, Spanish and German. Elsewhere, English, French, and Spanish have been the focus of contrastive studies with the aim to add to the

² KIAP refers to Kulturell Identitet i Akademisk Prosa which in English translates as “Cultural Identity in Academic Writing”

³ specon refers to Spanish economics papers

underdeveloped literature on academic discourse in French and Spanish (Carter-Thomas and Jacques 2017). As this PhD demonstrates in Chapter 5 and Chapter 6, the comparison of English to two major world languages shows that, although French and Spanish both share roles as major world languages and appear to differ from English in similar ways (Bennett and Muresan 2016), they do not always differ to the same degree or act similarly to one another. Furthermore, both French and Spanish were chosen as they are global languages on which limited study has been conducted in the context of academic discourse (Salager-Meyer 2000 p.26). Finally, these languages are suitable due to my own familiarity with the languages.

Similarly, economics was chosen for a number of reasons. First, previous research on KIAP has indicated that economics demonstrates interesting examples of metatext that is not always as prevalent in linguistics or medicine (Dahl 2004). Second, research on economics writing is quite limited. Overall, the corpus is a small, specialised corpus, composed of 150 research articles with 50 in each language, all in economics. Underpinning the development of KIAP-EEFS are strong principles of sampling and representativeness; the corpus has been well-constructed and delimits its focus to research articles that are captured within the same sampling frame. Therefore, KIAP-EEFS is only representative of economics research articles in English, French, and Spanish between the years of 1995 and 2003. The construction of KIAP-EEFS satisfies the principles of a comparable corpus, detailed in Section 3.5.4, which therefore ensures comparability and a dependable *tertium comparationis* at each stratum of data.

While this brief discussion of contrastive and corpus linguistics is only a cursory overview, the discussion in Chapter 3 and the presentation of the methodology in Chapter 4 give greater detail on each of the concepts discussed here. Overall, research on corpus-based contrastive analysis of academic writing is notably experiencing a rebirth in interest due to its role in a world of increasing ‘interlingual and intercultural communication’ (Granger 2003, p.18). This rebirth is largely influenced by advances in corpus linguistics over the last 30 years and, while interest in contrastive linguistics is growing, it largely remains restricted to the study of two languages. For example, in their 2017 edited volume, *Contrastive Analysis of Discourse-pragmatic Aspects of Linguistic Genres*, Aijmer and Lewis open by saying that ‘the domain of contrastive linguistics centres on the

comparison, in synchrony, of two languages’ (2017, p.1). Therefore, while this thesis profits greatly from the adoption of contrastive and corpus methodologies, an important contribution that this research makes to this growing field is its response to calls from the literature which see the value in comparing three languages and recognises that a third language can tell us something valuable about the other two (Van der Auwera 2012, p.75). Bearing this brief discussion in mind, the following section, Section 1.4, presents the research questions and delimitations that have guided and focused this study.

1.4 Research questions and delimitations

This section outlines the research questions and delimitations of this PhD thesis. Starting with the research questions in Section 1.4.1, the overarching questions that guide this thesis are presented. This is followed by consideration of other relevant questions that emerge surrounding wider issues in the literature. Following that, the delimitations that demark the focus of this study are discussed in Section 1.4.2.

1.4.1 Research Questions

As demonstrated in Section 1.3, this PhD research makes original contributions to the fields of academic discourse and corpus-based contrastive linguistics. These gaps pertain to the study of academic discourse in languages other than English, namely French and Spanish; the study of questions as devices of reader engagement which pertains to the field of metadiscourse in academic writing; the study of the economics disciplinary research article; and the corpus-based contrastive analysis of three languages, namely English, French, and Spanish. In addressing these gaps, this thesis answers two main research questions:

1. To what extent does the presence of questions as reader engagement correspond in English, French, and Spanish economics research articles?
2. To what extent does question function as reader engagement correspond in English, French, and Spanish economics research articles?

In moving to answer these research questions, other relevant sub-questions emerged that were necessary to answer in order to situate the context of this research. These questions allowed this study on questions in economics academic writing in English, French, and Spanish research articles to comment on wider themes and contribute to the understanding of current problematic issues in the literature. These questions follow:

1. What is the relationship between academic writing in English, French, and Spanish and how does this relationship impact the comparability of academic writing across the three languages?
2. What constitutes the discipline of economics in this study? Is it the same discipline in each language or does economics refer to something different in English, French, and Spanish?
3. What is the nature of the discourse community studied herein? Is it one discourse community spanning English, French, and Spanish or does this study analyse three separate discourse communities in English, French, and Spanish? What is the impact of these conceptualisations of discourse community on the comparability of academic writing across the three languages?
4. To what extent is the genre of the research article comparable across English, French, and Spanish both textually and contextually? How do theories of genre identify the research article? Is the economics research article a disciplinary genre and is it comparable in English, French, and Spanish?
5. What theories of comparability can ensure faithful comparisons in contrastive linguistics? What are the roles of the *tertium comparationis* and equivalence in contrastive studies? How do these concepts inform the approach to contrastive analysis employed herein?
6. What value does corpus linguistics offer to contrastive analysis? What research exists that merges these two fields? How do the findings of this study's corpus-based contrastive analysis relate to the current state of the art?
7. What differences and similarities occur in the presence of questions in terms of frequency, function, question length, question type and form, question location, use of passivity in questions, use of tense and verbal modality in questions, and question sentence types in English, French, and Spanish?
8. What differences and similarities occur in question functions in terms of frequency, question length, question type and form, question location, use of passivity in questions, use of tense and verbal modality in questions, and question sentence types in English, French, and Spanish?

The above questions have been used to guide the discussion of the literature and methodology in Chapter 2, Chapter 3, and Chapter 4, while the two main research questions frame Chapter 5 and Chapter 6, which present the results and findings of the analyses. These findings are discussed in Chapter 7.

1.5 Delimitations

In order to further narrow the scope of the study, it is delimited in a number of ways. Firstly, the analysis takes a convergent contrastive view, whereby items perceived to function similarly are contrasted, as opposed to a divergent approach where, instead of

contrasting different items in each language, a linguistic feature is identified in one language and the analysis seeks to find correspondences in other languages. A convergent approach was chosen because questions form the basis of the research and this research was interested in how questions compare across languages. However, taking a divergent approach would be an interesting next step in this research to assess whether questions best correspond to questions in each language in terms of reader engagement, or if they correspond better to another means of reader engagement. Unfortunately, it is beyond the scope of this research to present such an analysis. Secondly, only such questions that occurred within the confines of the research article and that constitute part of the authorial text are analysed *i.e.* questions that occurred within references and tables were excluded from the study. Thirdly, only research articles from the discipline of economics in English, French and Spanish were studied. Fourthly, while this thesis is concerned with research articles, it has not discussed the impact of the publication process on the language of the research article. While it is interesting to consider the types of feedback and revisions suggested to writers and the impact this can have on published papers, the research articles studied in this thesis are seen as a communicative event that is not necessarily representative of one (or more) author's writing practices. Rather, they are seen as the accepted standard of a discourse community. This includes the impact of editors and reviewers on the writing, which raise it to an international standard of academic scientific writing, and not that of individual authors (Fløttum *et al.* 2006, p.11). Finally, questions are only analysed in terms of frequency, function, question length, question type and form, question location, use of passivity in questions, use of tense and verbal modality in questions, and question sentence types. Further analyses in terms of negation, verb typology *etc.* could be considered; however, this remains an opportunity for future studies.

1.6 Chapter outline

This thesis is composed of eight chapters. Before proceeding to Chapter 2, each chapter is now outlined in detail.

Chapter 2, the first of two literature reviews, expands upon the context and rationale discussed briefly in Sections 1.3.1-1.3.2. Focussing directly within the fields of academic discourse, reader engagement, and the specific marker of questions to which this work

contributes, this chapter briefly situates this study within work on academic discourse in each language. Starting broadly, this chapter quickly zones in on the focus of academic writing and demonstrates the importance of discipline and the discourse community in conceptualising the population being studied *i.e.* expert writers in the discipline of economics. Consideration is also given to linguistic features of academic writing in the academic research article in English, French, and Spanish and, owing to limitations of space, the focus is specifically on the linguistic features that act as equivalences in this study *i.e.* structure, sentence length and type, tense and verbal modality, and passivity. The discussion of these linguistic features signals the direct location in the wider literature on academic discourse in English, French, and Spanish to which this PhD contributes. Following that, metadiscourse is discussed, with a focus on how interpersonal communication is achieved through dialogic writing. Metadiscourse refers to evaluation and, specifically, reader engagement in English, French and Spanish. Finally, the device of reader engagement studied in this thesis, questions, is discussed in detail, outlining the current state of the art regarding knowledge on questions in academic writing. This chapter is important for situating the research within academic discourse and for understanding the current state of knowledge on questions as devices of reader engagement in each language. The contribution this research makes adds new knowledge to each of these fields. The chapter closes by indicating the value of studies on languages other than English and links to the following chapter's focus on contrastive studies.

Chapter 3 presents the second literature review in this thesis, and builds on the brief discussion presented in Section 1.3.3. This chapter considers the theoretical challenges of comparing academic writing across languages, as well as the effective ways in which languages can be compared through a corpus-based contrastive linguistic paradigm. Building on the value attributed to contrastive studies, this chapter outlines the theoretical underpinning of contrastive analysis in order to make clear exactly what it is and that for which it is used. This chapter also delves deeper into the discussion of concepts in contrastive analysis such as equivalence, correspondence, and the *tertium comparationis*. In order to make clear the relationship between contrastive linguistics and corpus linguistics, this chapter also looks at representativeness, sampling, and multilingual corpora in order to demonstrate the value contrastive research adds to corpus linguistics and to show how comparability is ensured in the data analysed in this thesis. The chapter

closes by presenting research to-date that contrasts academic writing across languages under a corpus-based contrastive methodology, and what these studies say about academic writing, metadiscourse and reader engagement across the languages. This chapter is important for understanding the methodology used in this study, the relationship between corpus and contrastive linguistics, and how contrastive studies of academic discourse can add to our understanding of reader engagement across languages.

Chapter 4 presents the methodology of this thesis and discusses the data and questions studied herein, the *tertia comparationis* and equivalences used herein, and presents an outline of the two contrastive analyses presented in this thesis. First the data are discussed, focussing on the compilation of the KIAP-EEFS corpus, and then the process of extracting direct and indirect questions from the corpus is presented. This details the search processes used to find questions in the corpus and explains how direct and indirect questions are not always easily corpus-searchable. Following the presentation of the data, the analyses are then considered, beginning with an identification of the *tertia comparationis* and the equivalences used herein. Finally, the limitations of the study are presented before concluding the chapter.

Chapter 5 is guided by the first research question listed in Section 1.4 and presents the first of two corpus-based contrastive analyses. The focus of Chapter 5 is on the overall patterns in the use of questions as reader engagement in each language. These results offer a clear understanding of what English questions are generally like, what French questions are generally like, and what Spanish questions are generally like in terms of frequency; question length; question function; question type and form; indirect questions' illocutionary force indicating devices; location; passivity; tense and verbal modality; and sentence type in the context of these data. This is an important benchmark as these findings are then compared to those in Chapter 6 in order to assess the extent to which functions behave differently from the overall patterns of use in each language. In concluding this chapter, the manifest similarities and the overall correspondences across languages are presented.

Chapter 6 is guided by the second research question listed in Section 1.4 and presents the second of two corpus-based contrastive analyses, this time focussing on question functions as reader engagement in each language. The results of the analysis of each of

the following seven functions are presented: getting attention and focussing the reader, framing the discourse, organising the text, creating a niche, expressing an attitude and counter-claiming, setting up claims and protecting the writer, and asking real questions and pointing forward. These question functions are derived from the modification of Hyland's framework of question functions (2002a). Each question is categorised according to one of these seven functions and the questions in these functions are then tested in terms of the specified equivalences *i.e.* frequency; function; question length; question type and form; indirect questions' illocutionary force indicating devices; location; passivity; tense and verbal modality, and sentence type. It is of note that, owing to the complex nature of this chapter, which deals with a number of variables across seven functions in three languages, this chapter contains many short tables. However, they do contain pertinent information. In essence, this chapter shows how questions are used in each function in each language. This is important as it will make clear whether any functions behave differently across language and in comparison to the overall patterns of use in each language, outlined in Chapter 5.

Chapter 7 is the penultimate chapter in this thesis and presents the discussion of the findings presented in Chapter 5 and Chapter 6. This chapter makes ostensive links between the findings of this research and relevant facets of the literature to which it contributes. First, this chapter offers a discussion of the new insights gleaned from this study in the context of academic metadiscourse, reader engagement and languages other than English. This begins by discussing the findings from the analysis of the two *tertia comparationis* followed by a detailed discussion of each of the equivalences, reflecting on literature on English, French, and Spanish academic writing. Following that, the contributions to the multilingual academy and corpus-based contrastive linguistics are discussed.

Finally, Chapter 8 concludes this thesis by returning to the contributions, research questions and wider themes presented in Chapter 1, in order to summarise the findings and original contributions of this research. The contributions of this work are considered in terms the implications for the teaching and learning of academic language. Finally, this thesis concludes by looking forward to potential follow-up research projects that could address some of the questions that have emerged from this PhD research.

Chapter Two

Academic writing, reader
engagement, and questions in
English, French, and Spanish

2 Academic writing, reader engagement, and questions in English, French, and Spanish

2.1 Introduction

This chapter centres on research on academic writing in the academic research article in English, French, and Spanish. First, in Section 2.2, the discussion centres on the role of English as a lingua franca of academia, its impact on other academic languages, and the role of academic traditions and cultural identity on academic languages. The goal of this section is to situate this study within research on the global position of English and other academic languages as well as to better understand language as a variable in academic writing research. In order to identify the important gaps that this research addresses in the context of research on academic discourse in languages other than English, the focus is largely on issues of comparability which problematise the comparison of academic languages other than English with academic English. Next, to address the questions raised in Section 2.2 surrounding comparability, the disciplinary discourse community is discussed in Section 2.3. Focusing first on discipline, this section endeavours to demystify the concept by recognising the role of discipline as a social community. Following that, it moves on to discuss the concept of a disciplinary discourse community. Owing to the multilingual nature of this study, this section further problematises the role that language plays in defining whether this thesis deals with three separate discourse communities or one discourse community spanning English, French, and Spanish. Then, with the discourse community established, the shared genre is considered in Section 2.4. The goal is to identify the meaning and role of genre for this study and to establish the research article as a disciplinary genre that is comparable across languages in terms of both its textual properties and social and discursive context. This leads to a discussion of the textual features of the research article in English in Section 2.5, in French in Section 2.6, and in Spanish in Section 2.7. Each of these sections considers the structure of the research article, the length of sentences and sentence types used therein, and the role of passivity, tense, and verbal modality in research articles in each language respectively. As there is a dearth of research on the economics research article, these sections draw, where appropriate, both from relevant research on the economics research article and from the wider collection of studies on academic writing in English, French, and Spanish.

The purpose of these sections is to give insight into the linguistic features of academic research articles in each language, and to explicitly focus on the features that are used to analyse questions in Chapter 5 and 6. In Section 2.8, the contextual elements of the research article are considered with an explicit focus on the interactive and social nature of the genre, looking broadly at metadiscourse. The section then moves towards a focus on evaluation in academic writing and the roles of writer stance and reader engagement, which are discussed in Section 2.9. The goal of Section 2.9 is to outline the current state of research on evaluation, writer stance, and reader engagement, signalling clearly that reader engagement, which refers to the ways in which writers position and exploit readers within texts, is relatively underrepresented in the literature in English, and especially in French and Spanish. Next, Section 2.10 focuses explicitly on the use of questions in academic writing as a device of reader engagement. It then synthesises the existing body of literature on questions in order to delineate the framework of questions as reader engagement, question functions as reader engagement, and question types and forms. This section also makes clear that this is an under-researched area in applied linguistics and academic discourse research. Finally, Section 2.11 concludes this chapter, summarising the main points and gaps identified.

2.2 English academic discourse and the multilingual context

As discussed in Chapter 1, Section 1.3, the field of academic discourse broadly refers to the study of the language and ways of thinking within academia. As a field of enquiry, academic discourse has received much attention in the social sciences, with the most substantial body of work belonging to applied linguistics (Heller and Morek 2015, p.174). Owing to the global presence of academic institutions, research and descriptions in academic discourse envelop a range of dynamic, flexible, and complex discourses (Coffin and Donohue 2012, p.71; Heller and Morek 2015, p.179) which vary across languages, cultures, genres, and modes (Flowerdew and Wang 2015, p.82). This diversity of linguistic description has led to research on academic discourse informing the teaching of academic language (Bhatia 2002a, p.21). However, as discussed in Section 1.3, while it is hoped that the findings of this thesis will be of use to those teaching academic language⁴, this thesis is primarily concerned with delivering a description of academic writing. This

⁴ For more on the application of this study to academic language pedagogy, see the detailed discussion in Chapter 8, Section 8.3.

description is envisaged to help us better understand the roles, identities, and practices of members of the academy, in the context of the language of the economics academic research article in English, French, and Spanish.

The multilingual nature of this study poses a number of challenges, particularly with regards to the notion of comparability, which is discussed in detail in Section 3.3.1. Essentially, ensuring comparability is necessary when conducting contrastive studies, since it is imperative that like with like is compared in order to produce faithful results. While this may appear an obvious prerequisite, the challenge posed to this study is not just of ensuring that the economics research article is comparable in English, French, and Spanish – although this is also important and is discussed later in this chapter in Section 2.4. Rather, the primary challenge pertains to ensuring comparability when comparing any language and domain of language use with English. This is a challenge recognised by Swales (2004, pp.52-53), who quite appropriately argues that comparing languages other than English with English can be intrinsically problematic due to the sheer vastness of speakers and users of English as well as its global standing and international prestige that can render English incomparable to other languages.

The impact of English as the dominant academic language on other academic languages forms part of a large debate in critical English for Academic purposes and academic discourse studies (Benesch 1996; 2001; Pennycook 1997; 2001). To focus on relevant research, the discussion briefly covers two main perspectives. First, in Section 2.2.1, the relationships between English and other languages of scientific and academic communication are discussed with a focus on the growth of English as the lingua franca of academia and the impact that this growth has had on other languages; where possible, specific attention is paid to French and Spanish, as they are the languages studied alongside English in this thesis. This discussion illustrates the complex landscape of academic discourse globally; outlines the challenges of comparability when comparing languages of different global standing; and signals the value of pursuing research on academic language in French and Spanish, in the field of academic discourse. Second, in Section 2.2.2., cultural identity in academic writing is discussed, highlighting how academic traditions and cultural identity shape the linguistic features of academic discourse produced in different parts of the world. This shows that different cultural

backgrounds result in varied approaches to producing academic knowledge, rendering language an important variable across academic discourse, and further problematising the assumption of the comparability of academic discourse in a multilingual context.

2.2.1 English as the lingua franca of academia: the impact on the multilingual academic context

To-date, most studies on academic discourse have focused on the English language, due to its ‘privileged position [...] in the world of science today’ (Fløttum *et al.* 2006, p.18). This is a position that it has readily occupied since its replacement of German post-1914 as the international language of science and research (Ferguson 2007, p.12). The increased presence of English in academic contexts has led to important transitions in the field and the academy, mentioned briefly in Section 1.3. For example, English’s presence has grown owing to the greater emphasis placed on writing in English in higher education (Klein and Boscolo 2016, p.311), the accelerated need for the teaching and learning of academic English (Dewey 2007, p.333; Seidlhofer 2010, p.366), and the importance of English academia and scholarship to the global construction of knowledge (Hyland 2009, p.185). The first two points relate more closely to pedagogical research. Such research seeks to support the teaching of academic language to those studying in English speaking countries and English-medium instruction contexts where, in countries like Spain, government mandates require that lecturing staff deliver some courses in English (Dafouz and Camacho-Miñano 2016, p.58). While this is of course an interesting phenomenon that furthers the reach of English over academic domains, this thesis is more concerned with the third point on the importance of the English academic language to the global construction of knowledge.

With English as the lingua franca of the academy, ways of thinking, constructing, and sharing knowledge inherent in English-language academic culture have been privileged and academic writing in English has been seen as a rule setter for other languages (Connor 1996, p.31). The epistemological traditions of English as academic discourse vary greatly from those in French and Spanish where, following Bennett (2015), the ‘expansion of English academic discourse (EAD) as the only acceptable vehicle for knowledge in the modern world’ has resulted in the ‘imposition of the empiricist paradigm as the only valid form of knowledge in the modern world’, which occurs very much at the cost of

‘continental philosophy’ and other theories of knowledge (pp.11-14). The empiricist paradigm is so recognised that it has been awarded not only the highest status in knowledge construction but also a position of objectivity and infallibility that is arguably incomparable when compared to ‘German idealism, hermeneutics, phenomenology, [and] French poststructuralism’ (Bennett 2015, p.14), for example.

The growth in empiricist values of knowledge has also positioned the research article, a genre which has been dubbed ‘the master narrative of our time’ (Montgomery 1996), as the primary means of disseminating knowledge. The primacy of the research article has created a “publish or perish” culture in the academy, where publishing peer-reviewed research articles is linked to career progress and funding (Tijdkink *et al.* 2013; Tijdkink *et al.* 2014). This gives rise to neoliberal and capitalist metrics in determining the quality of scholarly research (Phillipson 2006, p.15). ‘Publish or perish’ is a well-known phrase among the members of the academy, and may be better expressed as ‘publish in English or perish’ following Ammon (2006, p.11) and Koutsantoni (2007, p.47), or better still, publish in French or Spanish and perish (Schluer 2015, pp.237-238). Expectedly then, more and more studies are published in research articles in English year on year with publications in other languages in decline (Ferguson 2007, p.17). This phenomenon further defends the hegemonic status of the empiricist perspective as one of greater value. Moreover, this continues to position the English language research article as an incommensurable means of constructing and sharing knowledge.

While English language publications are generally readily available in online international journals, the same cannot be said for French and Spanish language publications. These are often published in traditional paper formats as opposed to online journals (Bennett 2015, p.14). French and Spanish academic traditions do not share the same ‘publish or perish’ culture that exists in English, as they have a much shorter history of publishing internationally and publishing in empiricist paradigms (Pérez-Llantada *et al.* 2011, p.22). As such, less pressure is placed on academics to publish in French and Spanish. This occurs for a number of reasons, and the primary cause, as has been well-argued elsewhere (Breeze 2015; Uzuner 2008), is the fact that publishing in any language other than English reduces reach and readership and can negatively impact career progress and research funding opportunities. This shift towards publishing in English has greatly impacted

French and Spanish knowledge cultures, as noted by Pérez-Llantada *et al.* (2011, p.22), who document that one of their subjects found the sudden need to publish internationally in English in the format of the research article to be very challenging:

In Spain there has been a tremendous change in the past 40 years. In the past, there was no perceived need for publishing. We lived in isolation for whatever reasons. Publishing was not common, and we only published in national journals produced by the Spanish Royal Academy. It was all at a local level, not because people were not qualified, but simply because people did not think about the fact of publishing.

As well as the research article, French and Spanish also contain a wider variety of acceptable means of academic communication such as constructivist, prose-like, and more literary discussions of research which are typically considered non-academic and less prestigious in English when compared to the research article (Bennett 2015, pp.12-13). Overall, research that is published in French and Spanish is often scattered and not widely disseminated (Carter-Thomas and Jacques 2017, pp.2-3). As such, many publications in French and Spanish remain invisible to the wider international English-speaking academic community; this is a problem that impacts this study too, since there were far fewer studies published in French and Spanish available to inform the literature review.

Due to the move towards non-anglophone scholars publishing almost exclusively in English, ‘non-Anglophone scholars [appear] linguistically disadvantaged relative to native speaking academics when it comes to publication’ (Pérez-Llantada *et al.* 2011, p.45). While researchers like Bocanegra-Valle (2014, p.1) argue that a main cause of this disadvantage is linguistic, others like Hyland (2016, p.67) refer to linguistic disadvantages as a myth, arguing that ‘the current orthodoxy which attributes publishing success to mother tongue perpetuates an idealized monolingualism that still underlies a lot of thinking in applied linguistics’. Nonetheless, at the very least the influence of first language on academic writing has largely been found to be non-negligible (Carter-Thomas and Rowley-Jolivet 2013, p.112). Already, the academic discourse of English, French, and Spanish appears complex and challenging to equate and an important question emerges regarding the impact of the position of English as the dominant global lingua franca of academia on French and Spanish academic language.

With the growth of English as the lingua franca of academia and science, there emerge important conversations on the impact this has on academic discourse in other languages. Academic English has been ambivalently accepted, globally. To some, it represents a

neutral middle ground which allows academics to connect with one another through a shared language on an international stage (Haarman and Holman 2001, p.256); to others, it is not necessarily a new phenomenon as there have always been international and national languages of science. Internationally, prior to English, this scientific domain was occupied by German (Ferguson 2007, p.16). For others, it is less a lingua franca and more a 'lingua frankensteinia' (Phillipson 2008a) or 'tyrannosaurus rex' (Swales 1997), which threatens other languages and ways of thinking in the academy (Bennett 2015, p.25). The view of English as a neutral lingua franca that simply serves international exchange typically rests upon the claim that as an academic discourse, and one strongly embedded in the empiricist paradigm, it is impersonal and objective in nature. This, of course, is quite a serious and misguided claim, and Section 2.9 focusses explicitly on how writers' expressions of subjectivity in academic writing through the use of rhetoric is well established. From a critical perspective of academic discourse, it is important to recognise the role of power inherent in the use of English as the lingua franca of academia, which can result in the disadvantaging of second language speakers of English for publishing, as previously discussed in this section, as well as facilitate domain loss and diglossia in other language contexts, concepts which are now discussed.

Domain loss is a phenomenon in language study where a domain or context in which a language has traditionally been used is now occupied by another language. For example, domain loss could be the use of English in academic writing or teaching where, in the past, national languages would have held primacy. As Mufwene (2005, p.24) notes, languages that share functions and space tend to be in direct competition with one another, and if academic English and another academic language share a space, English has traditionally emerged the victor. The result of such a domain loss is wide-ranging and can result in information loss, linguistic atrophy, and academic deskilling, for example. Information loss refers to the likelihood of losing ways of thinking and constructing knowledge, known as "epistemicide", or the product of that knowledge where ideas and solutions produced in other languages do not receive international attention nor add to the international knowledge basis. In terms of epistemicide, the previous discussion in this section on the imposition of the English language empiricist paradigm on other languages at the cost of other ways of thinking is of note. Bennett (2015, p.11) discusses a startling vision of the future of the academy where, following the silencing of and decrease in

publishing in other academic languages, there is a risk of a monoculture of thought. Focusing on the product of knowledge, Breeze (2015, p.41) summarises the risks of information loss through a lack of engagement with research in languages other than English by referring to Swales and Feak's (2000) account:

of a crano-mandibular muscle first fully described in a publication in German in 1954, which was then "discovered" by Portuguese scientists in 1978, and finally "discovered" for the English-speaking world in 1996. Each group of scientists presented this finding as entirely original, we assume, in good faith, because language barriers made the work of the earlier group(s) inaccessible.

Information loss risks slowing scientific progress through the exclusion of ideas and discoveries not published in English. For example, studies that focus on the analysis of culture are typically written and published in the language of that culture (Bennett 2015, p.13) and, as such, the erasure or relegation of an academic language to a lesser stage may impact the way research on other language contexts and cultures is discussed and presented to the world. Coupling this with the development of a monoculture of thought which excludes means of thinking that exist beyond the empiricist epistemological paradigm, the domain loss of languages other than English to English in the academy is problematic. Further challenges that languages face lie in academic deskilling and linguistic atrophy, with languages losing their ability to engage with and discuss research (Phillipson 2006; Ferguson 2007). Owing to such consequences of domain loss, academic English can be viewed as a threat to academic writing in other languages and can create diglossia in academic discourse where academic languages other than English are seen as less valuable, less prestigious, and less impactful.

The relegation of an academic language other than English to more parochial roles in education and research has been discussed at length and is summarised by Ferguson (2007) and Uzunur (2008). A core argument therein pertains to the reality that the international role of English as a lingua franca of academia does not necessarily eliminate nationally disseminated research which remains of value to educators, government, and practitioners in their relevant languages and contexts. Furthermore, in the case of French and Spanish academic research, there is clearly a well-established body of knowledge published in peer-reviewed journals (Breeze 2015, p.41). However, to accept uncritically the view that these factors indicate that French and Spanish's academic languages peacefully co-exist with English would be negligent, as there is ample evidence of both a

decrease in publications in languages other than English, including French and Spanish, and an increase in publications from non-English speaking countries in English – see Ferguson (2007). These decreases are often due to policy interventions, such as the Bologna 1999 agreement. This agreement has achieved great ends in the development of the European Higher Education Area, but must also be held to account for its hegemonic gatekeeping of the role of English at the centre of international collaboration and research, which it does in order to support the endeavour to contribute to and generate a common knowledge (Phillipson 2006; 2008b). Overall, while the views that the national language still serves national value in academic contexts are accurate, this does not dispute the diglossia that the relegation of academic language creates. Though it is argued that there have always been international languages of academia, the neoliberal landscape of academic publishing and the value systems and scientific metrics for determining quality in research have changed the academy. This change has fostered the publish or perish lifestyle within the academy, which determines career progress and funding and, as such, renders the current face of academic publishing incomparable to a former lingua franca of academia such as German, as there is an evident decline in publishing in other languages unlike ever seen before. Furthermore, the relegation of languages to a national level indicates that readership of research published in French or Spanish is not the same as that of English. Considering the importance of concepts like ‘audience’ and ‘communicative purpose’ (Askerhave and Swales 2001, p.209) to academic discourse communities, a concept which is discussed in detail in Section 2.3, the unequal position, purpose, and readership of English, French, and Spanish academic discourse calls into question the possibility of faithfully comparing academic writing across languages.

The concerns discussed in this section are important for this study in a number of ways. First, they show that the terrain of academic discourse is both global and local and is not one unclouded by issues of power and privilege. Moreover, the concerns raised above challenge this study to define the most effective means to ensure comparability when comparing across languages. However, more than simply a means of guiding comparability, these concerns draw attention to important and often underrepresented issues in applied linguistic research on academic writing. That is, owing to the domain loss and diglossia that relegate languages other than English to parochial roles in education and research, there is a potential for information loss, linguistic atrophy, academic

deskilling, and changes in readership in languages like French and Spanish. Overall, there are much fewer papers produced in these languages when compared to English, and their visible production is in decline. As a result, there are fewer studies on academic discourse in French and Spanish, at least in online academic journals. This study, recognising this important gap, contributes to research on academic writing in French and Spanish as well as English, in order to add to the body of knowledge that recognises a value in better understanding academic discourse in languages other than English and the value academic discourse in languages other than English offers to the wider academy.

Overall, the broad perspective discussed here helps us to better understand the relationships between academic discourse in English, French, and Spanish and the roles that lingua franca status, domain loss, and diglossia play in its development and evolution. Touched on briefly above was the concept of epistemicide - the idea of the eradication of ways of thinking owing to domain loss. Epistemicide presupposes the idea that different languages construct knowledge differently, which implies there is an influence from cultural concepts like academic tradition and cultural identity on academic writing in different languages. In the next section, the role of culture is discussed with a view to illustrate the differences between academic language in different languages and to demonstrate the importance of language as a variable in academic discourse studies.

2.2.2 Academic traditions and cultural identity: the impact on academic texts across languages

In section 2.2.1, the first discussion considered the conflicts between the roles of English and languages other than English in the academic community. In this section, the discussion demonstrates how academic traditions and cultural identity can shape the nature of academic writing in different languages, which accounts for a further variable when comparing across languages, namely culture. In fact, following Duszak's (1997, p.11) work on variation in academic discourse, it is clear that culture impacts greatly on academic writing, especially in the humanities and social sciences:

Among the most notable differences are field- and culture-bound disparities in global organisation of schemata of texts. These include divisions within text space, their labelling and sequencing. Texts have been found to vary in the degree of explicitness and metadiscoursal guidance as to what meanings have been, are, or will be communicated. They have been shown to differ in redundancy levels and in the amount of background that is sanctioned in establishing relations between ideas. Differences in the use of structural resources and rhetorical devices have also been pointed out.

To add to this view, it is important to consider Connor (2002, p.497) who makes the point that ‘readers’ expectations determine what is perceived as coherent, straightforward writing’. That means that it is important to be cognisant of the evaluative language that has been used in the literature for describing different languages *e.g.* “clear” or “easy to read”. These terms are culturally bound, culturally relative, and must not be understood as a judgement of the quality of academic writing in a different language. In fact, Spack (1997) addresses this concern in her criticism of cultural insensitivities disseminated within contrastive rhetoric descriptions. Therein, she identifies that the designation of a language as reader-friendly is entirely culturally relative. That is to say, while it is possible to claim that French or German writing is generally more difficult to follow and read for Anglophone readers, this does not mean that the Francophonie or Germanophonie find it equally challenging. To argue such a view would be Anglocentric, lending equivalence between reader-friendliness and English language norms. However, as this project is not an investigation of culture, but rather uses the wider cultural contexts and identities as one means of interpreting the results, it is beyond the scope to address these criticisms. Key researchers in the field such as Mauranen (1993), Connor (1996; 2002; 2003), Salager-Meyer *et al.* (2003), Connor *et al.* (2008) and Fløttum (2012) have discussed these concerns. What is most important for this research is to understand how traditions and culture impact academic writing in different contexts.

Across languages, it has been shown that academic tradition and culture impact the nature of academic writing. Clyne (1994, p.238) argues that German academic writing idealises knowledge, whereas English favours linearity and relevance of information. Nichols (1988, p.400) presents Russian academic writing as communicating fact as opposed to a persuasive rhetoric. Mauranen (1993, p.4) identifies Finnish as impersonal, Duszak (1997, p.14) positions Polish writers as defensive, and Cmejrkova and Danes (1997, p.42) identify Czech academic writing as not well arranged or systematic. In terms of Spanish cultural impact, Orta *et al.* (2006, p.199) find that Spanish academic writing tends to show less stance, while Salager-Meyer *et al.* (2003, p.223) find that French and Spanish have traditionally been more critical in academic writing than English, and in book reviews Spanish is more critical than English and French (Salager-Meyer and Ariza 2004, p.156). Similarly, Bachschmidt (1999, p.27) finds that French tends to focus on new information and facts instead of focusing on persuasion and existing information. Fløttum *et al.* (2006,

p.191) argue that English is more explicit in nature and French is less explicit, and Salager-Meyer (2011a, p.36) argues that research papers in French are more prescriptive and authoritarian and that research papers in English tone down their claims to avoid face threatening acts – a feature that she argues is the most prominent cultural difference between English and French academic writing. She positions French academic writing as being exaggeratedly self-confident. Moreover, Salager-Meyer *et al.* (2003, p.223) identify that French and Spanish academic writing tends to be more passionate while, they argue, the opposite is true for English. Kruse (2013, p.40) identifies that French academic writing is largely influenced by four modes of texts, which Donahue (2002, p.136) labels ‘*commentaire composé*’, ‘*dissertation*’, ‘*étude d’un texte argumentative*’, and ‘*discussion*’, while for Spanish, Kruse (2013, p.40) also identifies that intellectual thought patterns were institutionalised following the move towards national academic languages from Latin. Therein, she finds that there is an absence of hedging, or modalisation which she argues are traces of Spanish culture in academic writing. Similarly, research on politeness in academic writing has revealed cultural differences across language, and Lafuente-Millán (2014, p.208) identifies that:

members of Anglo-Saxon cultures generally prefer negative politeness strategies which reflect their values, such as autonomy of the individual and respect towards the right of others to make their own choices. In contrast, the Spanish tend to favour involvement strategies addressed at the positive face of their audience.

Overall, the range of cultural differences across English, French, and Spanish and their impact on writing is extensive and well-established, and, in summary, Bennett and Muresan (2016, p.114) have identified that:

there exists an approach to scholarly writing in the Romance cultures that is so different from [English] that it could be considered a distinct discourse. Unlike English, it does not aim to transmit the ‘facts’ as concisely and straightforwardly as possible, but rather delights in linguistic complexity and nuance, presenting all data as filtered through the consciousness of the human observer. In this regard, it could be said to encode a different theory of knowledge to the empiricism that underpins the English plain style, a humanistic orientation that is more holistic, in that it values subjectivity (aesthetics, ethics and emotivity) as essential components of knowledge.

Many of these differences relate to research on metadiscourse, which is discussed in greater detail in Section 2.8. Overall, while it is not possible to identify every way in which culture impacts academic writing, what this research provides is proof that it can. In terms of categorising or labelling language cultures, the concepts of reader-responsible and writer-responsible languages, and content-oriented and formal-oriented languages,

have been important in the literature when considering the interactional nature of writing, as this thesis does. Therefore, these concepts are now discussed in more detail.

Following on from Kaplan (1966), Hinds (1987) uses the terms reader- and writer-responsible in his typological analysis of Japanese and English, where he categorises Japanese as reader-responsible and English as writer-responsible, owing to differences in clarity and coherence on the part of the writer. In his view, reader-responsible languages place the onus on readers to navigate and understand texts, while writer-responsible languages place more responsibility on the writer to ensure that readers understand texts. This view can be applied to other languages; for example, Clyne finds that first language academic writers of German place the responsibility on the reader to effectively navigate a text, constructing a reader-responsible context (1987, p.238). Of particular note is research by Salager-Meyer (2011b, p.71) on English, French, and Spanish which outlines the nature of reader- and writer-responsible languages as follows:

Contrastive rhetoric analyses of scientific discourse have drawn attention to the existence of differences in the level of explicitness between languages. English is usually said to lie at the higher end of the scale of explicitness of text organisation, clarity, and coherence: English readers indeed expect and require landmarks of coherence and unity as they read, and writers need to provide these transitional statements. Texts written in English thus reflect a more reader-oriented attitude: explicit statements are regarded as polite to readers and implicitness as impolite. When compared with the scientific rhetoric of Anglo-American writers, writers in other languages are much less inclined to regard explicitness as their responsibility. [...] Research articles written in Spanish and texts written in Portuguese are also negatively marked as to the presence of cohesive order to indicate the discursive logic of texts. A much lower density of periphrastic links, previews, and reviews has also been noted in scientific papers written in French and in Slovene and German academic writing.

In this view, English is writer-responsible while French and Spanish are considered reader-responsible. This perspective is supported by others such as Mur-Dueñas (2007a, pp.46-47), whose study of business management articles in English and Spanish finds that English uses more contrastive logical markers such as “however” than Spanish. Mur-Dueñas argues that this contributes to the view of English as being more writer-responsible and Spanish more reader-responsible. Bachschmidt (1999, p.26) identifies that French exhibits more reader-responsible behaviours; there is evidence of much fewer links between arguments in French academic writing than in English. This is found to impact the overall readability of the text.

Building on this work, a similar concept proposed by Clyne (1994) considers languages as being of content-oriented or formal-oriented cultures. Adopting Clyne's work (1994), Cuenca (2003, p.1085) and Oliver del Olmo (2004, pp.112-115) present content-oriented languages as those that tend to place responsibility on the reader, and value the quantity of knowledge presented. This differs from formal-oriented cultures, which are writer-responsible and are more concerned with the way knowledge is presented. These formal-oriented cultures tend to have deductive and linear texts, for which the opposite is true in content-oriented cultures. Content-oriented cultures often use complex language and are more indirect than formal-oriented cultures which typically use more concrete language. As such, formal-oriented cultures typically use short sentences, repetition, and content-oriented cultures use complex sentences and avoid such repetition. Overall, given the preference for linearity, formal-oriented cultures tend to produce systematic texts which are consistent in their organisation, while texts originating in content-oriented cultures differ and can appear more complex and varying.

Academic writing in English, according to Hyland (2005b, p.28) is explicit in nature, like formal-oriented cultures. Others like Van Bonn and Swales (2007, p.102) have found that French exhibits longer sentences, which are in keeping with content-oriented cultures. Mur-Dueñas (2011, p.3072) has found that, owing to the lack of linearity in Spanish, there is an increased need for topicalisers which accompany 'lengthier, more elaborate arguments'. Overall, her research positions Spanish as content-oriented and English as formal-oriented. While this area of research is not densely populated, it is consistent when considering the descriptive views of English, French, and Spanish academic writing. Overall, the view of English as reader-oriented and formal-oriented, and French and Spanish as writer-oriented and content-oriented is generally accepted in the literature (Hinds 1987; Clyne 1994; Valero-Garcés 1996; Cuenca 2003; Oliver del Olmo 2004; Pisanski-Peterlin 2005; Ädel 2006; Fløttum *et al.* 2006; Hyland and Salager-Meyer 2008; Amiryousefi and Rasekh 2010; Mur-Dueñas 2011; Salager-Meyer 2011).

The goal of these brief summaries of research is to explain some of the reasons why writing can differ across languages, specifically those that pertain to the reader-writer relationship, which are important for the discussion of results in Chapter 7. What is important to identify is that culture in academic writing evidently shapes the linguistic

features of academic discourse produced in different parts of the world and in different languages. Building on the first part of this discussion, which centred on the growth of English academic writing and domain loss in French and Spanish academic writing from a global perspective, the role of academic tradition and culture in writing further problematises the comparability of academic discourse in the multilingual context. This indicates that there are challenges not only without - in the complex global landscape of academic discourse - but also within, in the language used to perform this academic discourse being influenced by different academic traditions and cultures. While it is evident that there is a clear need for further research on languages other than English, there remains a question of whether it is possible to faithfully compare English to other languages, and whether it is possible to overcome the many obstacles discussed in this section. Traditionally, the field of academic discourse has resolved to study particular disciplinary discourse communities and specific genres in order to eliminate complicating variables in studies across languages; however, as the following section illustrates, grouping by disciplinary discourse communities and genres is not without its problems, with questions remaining as to the implication of language, location, and readership on the analysis of a disciplinary discourse community composed of speakers of different languages.

2.3 Disciplinary discourse communities

Recognising the challenges of working in a multilingual context identified in Section 2.2, the aim of this section is to better understand the role of discipline in academic studies, to identify the discipline of economics as a discipline in its own right, and to determine whether disciplines are comparable across languages as disciplinary discourse communities. First, discipline is considered with a view to categorise and substantiate economics as a discipline in its own right. This is done by conceptualising discipline as a social community, problematising membership of this community across languages, and by positioning the economics discipline that forms the basis of the data analysed in this thesis as constituting a focus disciplinary discourse community which transcends language barriers.

Following movements in the mid-twentieth century in applied linguistics, research on academic writing began to consider discipline as an important variable (Prior 1998,

p.196); a trend which has held strong today (Rongen Breivega *et al.* 2002, p.232; Salas 2015, p.20). However, what it is that constitutes a discipline is not exactly straightforward. For some it can be seen as an educational construction that departmentalises knowledge and learning (Seargeant 2012, p.115) while, in institutional contexts, discipline can be forged for non-taxonomic reasons (Mullally *et al.* 2016, p.23), where groupings are often guided by ‘marketing decisions, by political interests, or by administrative concerns’ (Thompson 2001b, p.15). Therefore, as a concept, it appears quite nuanced and reflexive.

The malleability and reflexivity of the concept of discipline is not surprising as the boundaries of what constitutes a discipline are quite blurry, especially when considering what constitutes a sub-discipline and how a sub-discipline differs from a discipline. Economics, for example, has consistently been analysed as a discipline in its own right, which comprises many sub-disciplines (Valero-Garcés 1996; Rongen Breivega *et al.* 2002; Dahl 2004; Moreno 2004; Fløttum *et al.* 2006; Carter-Thomas and Chambers 2012; Salas 2015). However, this identification appears to be relative, where sub-disciplines necessitate a broader discipline to bind them together, meaning that categorisation as a discipline or sub-discipline is largely based on the perspective of a study. That is, the decision to focus more broadly at discipline or more specifically at sub-discipline tends to depend upon the focus of the study, where more general conventions can be captured within “discipline” while more specific conventions within “sub-discipline” (Kanoksilapatham 2015, p.77). This view is important as a broad view of economics as a discipline allows for the generation of linguistic descriptions across the entire field; these descriptions could be studied in terms of sub-disciplines at a later date to test for variation, as Kanoksilapatham (2015) suggests. Moreover, the flexibility of defining a discipline or sub-discipline based on area of focus further fuzzies this polysemous term, rendering it rather tricky to define.

A further common categorisation of discipline takes a binary and often dialectical view of discipline as being either a “hard” or “soft” science. Hard tends to refer to scientific knowledge and natural sciences, while soft typically refers to humanities and social sciences (Becher 1994, p.152). Such groupings are somewhat crude and unhelpful given the complexity of knowledge and diverse approaches to its construction within these so-

called hard and soft disciplines. For example, the quantitative approaches of corpus linguistic research in the field have evolved to include complex statistical analyses, as well as the development of software that may have more in common with some “hard disciplines” than English literature, for example. In fact, the view of linguistics as a hard science has been argued in Yngve and Wasik (2006). Despite this, both corpus linguistics and English literature are typically seen as belonging the “soft disciplines”. In the crude hard/soft binary, economics has tended to be classified as a “soft discipline”, for example, Nagano (2010) identifies economics as such. While the field of economics includes many sub-disciplines with foci on mathematical and statistical formulae, the designation as a “soft discipline” here is not too problematic given the construction of KIAP-EEFS expressly avoided economics texts that contained many formulae or mathematical equations – the sampling frame and construction of KIAP-EEFS is discussed in Section 4.2.1.

In any event, it is clear that discipline is contentious, dynamic and by no means fixed, and there are evident overlaps in its conceptualisation (Darbellay 2015, p.170). This dynamic nature is not necessarily a negative characteristic, as research on academic writing has identified that disciplines can be studied in terms of interdisciplinarity (Klein 2014, p.1024). This looks at disciplines crossing over one another, and intradisciplinarity (Du Plessis *et al.* 2014, p.29), which is concerned with studies on variation within disciplines. Further categorisations have focused on multidisciplinary (Mobjörk 2010, p.867), which sees researchers from various disciplines converge to offer different analyses and perspectives on one item, and transdisciplinarity, which places discipline as a fuzzy continuum that has no ‘universal theory, methodology or definition’ (Klein 2013, p.189). These divisions reveal the growing importance placed on discipline in accounting for and creating differences in academic writing. Bearing all of this in mind, a number questions emerge. First, is it reasonable to say that economics is a discipline? Furthermore, is it a discipline that is comparable across languages?

Following Darbellay (2015, p.170) and myriad other studies mentioned earlier in this section (Valero-Garcés 1996; Rongen Breivega *et al.* 2002; Dahl 2004; Moreno 2004; Fløttum *et al.* 2006; Carter-Thomas and Chambers 2012; Salas 2015), the answer to the first question would appear to be yes. Economics is a discipline in its own right, which

encompasses many sub-disciplines such as applied economics and econometrics. However, given the problematisation of the comparability of academic writing in Section 2.2 and the evident fuzziness of the concept of discipline, is there sufficient evidence to ensure that economics as a discipline is comparable across languages? Does the discipline of economics contain the same knowledge, membership, and boundaries in English, French, and Spanish? It is difficult to say. That is because this study is not just concerned with the knowledge or the text, but also the academic social context in which a discipline is realised – a context, as made clear in Section 2.2.1, that is complex, international, and heavily nuanced.

From a social perspective, discipline is most commonly perceived as a social community or ‘academic tribe’ (Becher 1989, p.45) that shares norms, bodies of knowledge, conventions, and modes of inquiry and culture (Hyland 2004a, p.8). Taking the view of discipline as a social community lends a more tangible means of considering if the communities of economics writers in English, French, and Spanish, whose research articles are studied herein, form part of the same economics disciplinary community and are, therefore, comparable. In the literature, such a community, is usually referred to as a disciplinary discourse community.

The concept of a discourse community was originally applied by Kuhn (1970), who used the theory of community to analyse specialised communication among scientists. This provided an alternative theory of community to the likes of speech community, and became valuable for applied linguists studying academic discourse. As such, the concept evolved as others adapted it to their specific needs. For example, Bizzell (1992, p.222) saw it as ‘a group of people who share certain language-using practices’, while others such as Mauranen (1993) focused on discourse community in terms of the types of communities and contexts to which it could apply, *i.e.* scientific groups, academic groups, or social groups. However, most renowned among them is Swales and his seminal works on discourse community that arguably made the greatest advancements to the concept to-date (1990; 1993; 1998).

Swales spearheaded research on discourse community as a means to move beyond the constraints of speech communities. For him, discourse communities can be seen as ‘groups that have goals or purposes, and use communication to achieve these goals’

(Swales 1990, p.9). Moreover, the members of these communities must actively ascribe to them through conforming to discursive practices. With the focus on goals and purpose, there is not necessarily a geographical restriction on discourse communities which can limit alternative views of community, such as speech community (Borg 2003, p.398). For a discourse community, the core criteria for identifying the membership, according to Swales (1990, pp.24-27), state that members must:

- have communal and shared interests and goals;
- have mutual mediums for communication;
- make efforts to exchange information or knowledge;
- use genres that are specific to their community;
- use specialised lexis, terminology, language, acronyms; and
- be experts in the field or novices becoming part of the community.

A discourse community shares interests and goals (Swales 1988, p.212) and has shared ways in which to engage with these shared interests (Swales 1988, p.212). The community exploits these shared means of communication to realise goals and to exchange information (Swales 1988, p.212). This is achieved through genres, which are created and constrained by ‘discoursal expectations’ (Swales 1988, p.213) that ‘articulate the operations of the discourse community’ (Swales 1988, p.213). These practices also encapsulate ‘the development of community-specific’ (Swales 1988, p.213) terminology such as acronyms and abbreviations. To effectively grow the discourse community, there is a need for ‘a mass of members with a suitable degree of relevant discoursal and content expertise’ (Swales 1988, p.213). This conceptualisation of community has been typically used to analyse texts in the academic community, as in the case of this study. This is owing to the focus on language as creating the community (Price and Robinson 2016, p.5) and its express consideration of writing (Hyland 2015a, p.37) and, more specifically, academic writing (Swales 2016).

In general, discourse communities typically deal with written texts (Borg 2003, p.398), where membership is created by writer activity (Swales 1990, p.24). Therefore, with this in mind, is it reasonable to consider the writers of the research articles in KIAP-EEFS to

be part of a discourse community? Or do they constitute three separate discourse communities in English, French, and Spanish?

Considering that a discourse community is constituted by shared interests and goals, it is arguable that those writers publishing in peer-reviewed journals that inform the data of KIAP-EEFS share interests and goals in the conducting of research in the discipline of economics and sharing them in English, French, and Spanish. Moreover, following Sutton (1995, pp.38-39) the purposes of academic writing in the French context very much reflects this goal. Following the discussion of publishing practices in Section 2.2, the interest in so doing would be the dissemination of research for the purposes of knowledge creation or the development of an academic profile. The members of the discourse community are global too, owing to the topics of their publications in KIAP-EEFS which are not limited to national contexts in French and Spanish. For example, in the case of French, a journal that accounts for 68% of the articles in the KIAP French economics subcorpus, the *Annales d'Économie et de Statistique*, is defined as a ‘general purpose review covering the major fields of economic theory, econometrics and statistics. The review has taken a specific place recognized in the community of researchers both in France and abroad’ (JStor 2019). Moreover, for Spanish, the three journals used to construct the data play an international role with presence in Spain and Latin America – for details on the journals from which journal articles were extracted to build KIAP-EEFS see Sections 4.2.1 and 4.2.2.

With this in mind, it is reasonable to argue that the writers of the research articles in KIAP-EEFS share an interest in investigating and disseminating research on economics. Their main shared medium for communication is the research article; however, the research articles are produced in three languages that are not necessarily mutually intelligible to all members. This is an important consideration since, while a considerable number of researchers will be able to understand more than one of these languages, not all speakers will be able to directly communicate with one another. This means they are not necessarily exchanging information and knowledge across languages with one another. This is an important consideration as it questions the identification of the data studied herein as comprising one discourse community. This problem is considered in Section 2.3.1.

Returning to the operationalisation of this discourse community, the next consideration is shared language. Here, the research article, though not a genre that exists only within economics in each language, is indigenised to the community through shared linguistic practices (Fløttum *et al.* 2006; Lakić 2010; 2015; Salas 2015; Šćepanović 2015). Each language shares key terminology and jargon that feature in these articles and which are used to discuss the same key topics in a way that makes the economics research article sufficiently distinct from other disciplines like linguistics or medicine. Moreover, the speakers of each language identify the quality of the register of writing for academic disciplines as a formal and specialised register (Swales and Feak 1994; Cassany 1999; Salager-Meyer and Ariza 2004). Finally, in terms of membership, each language communicates with both experts and novices or new entrants. Although it may be difficult to measure this, the publication of these papers in peer-reviewed journals indicates the volition to communicate with peers in the field. Furthermore, many postgraduate students continue to forge careers in this area, under the supervision of experts, indicating that economics is a discipline that continues to enculture new members (Agence bibliographique de l'enseignement supérieur 2016a; Agence bibliographique de l'enseignement supérieur 2016b; CSUC 2016; Fundación Dialnet 2016; Higher Education Authority 2016; Ministerio de Educación 2016).

Overall, in many ways, economics disciplinary writers can be seen as creating one discourse community in three languages. Although, under this traditional definition by Swales (1990), that claim is somewhat tentative. In critiquing the concept of discourse community, it has been seen as idealised, inexact, and containing and being contained by fuzzy boundaries (Beaufort 1997, p.488; Hyland 2015b, p.25). For example, Swales sees a discourse community as not being bound by space and in so doing there is an implication that space and, by proxy, language and culture will not impact writing within the community (Prior 2003). However, Section 2.2 above clearly shows that language is an important variable that impacts both the social and global roles of academic languages as well as their textual manifestations. Given the multilingual focus here, the impact of language and geographical location on the discourse community is of particular interest. Breeze (2015, p.42) notes that the citation practices of English language publications typically exclude reference to languages other than English. Therefore, can these economists in English, French, and Spanish truly be part of a disciplinary discourse

community if in fact they do not talk to one another? The subsequent section moves to reconcile this problem by drawing on the concept of a focus discourse community.

2.3.1 Focus disciplinary discourse communities

Following critiques of his theory of discourse community, Swales (1998, p.22) recognised that his previous definitions of discourse community did not allow for multiple types of communities and was perhaps a little too idealistic. In order to overcome this, he proposed two subdivisions of discourse community, ‘place discourse community’ and ‘focus discourse community’ (Swales 1998, p.201).

Swales’ original conceptualisation of discourse community became place discourse community (1998, p.201). Place discourse community (Swales 1998, p.201) refers to a discourse community that is in physical contact and that shares practices, traditions, language, and norms (Morton 2016, p.54); therefore, for the reasons identified in Section 2.3, place discourse community would not prove an effective theory of community for the purposes of this research.

However, a focus discourse community is not constrained by concepts such as time, language, and geographical location (Prior 2003, p.2). It is named focus from it being a ‘grouping of people joined by a shared focus of interest’ (Thompson 2009, p.54). Much fewer studies have employed focus discourse community as a framework of analysis; however, using Swales’ well-known example (1998, p.201), a focus discourse community can be observed in the lives of three café owners who never meet, yet operate on the same busy street. Each one has never met the others, yet they all own a café, live/work on the same street, and engage in discourse practices such as ordering stock or menu making, *etc.* In this way, although they are not in direct contact, they are in contact by degrees of separation and part of this connection is created through their discursive practices. Prior (1998), who quickly took to this subdivision of discourse communities, identifies that these communities cannot be seen as stationary and that it is important to award them a degree of dynamism. This is especially true for instances where the subgroups of a focus discourse community can vary in size and composition. In the case of this study of English, French and Spanish, such a perspective addresses Swales’ concerns surrounding the comparison of languages other than English with English, owing to vastness of speakers of English (2004, pp.52-53).

These views of focus discourse community are very relevant to this research as the discourse community studied herein is composed of economists of English, French, and Spanish language backgrounds. As the writers belong to three different writer populations, it could be considered reasonable to see them as three separate disciplinary discourse communities, which would render a comparison more challenging. However, the academic discourse community is not monolingual and, although most research published today is published in English academic journals, for now there remain active and relevant academic journals in French and Spanish (Breeze 2015, p.41). Therefore, in moving to answer the questions posed in Section 2.3, this research proposes that each language represents an individual disciplinary discourse community and together they represent a focus disciplinary discourse community. Taking this view of focus discourse community, and although the writers in KIAP-EEFS may not share the bonds of space and language, across all three languages, these writers share goals in publishing and disseminating information; means of communication with one another; efforts in exchanging information; uses of genres specific to their community; language for discussing specialised topics, and expert members. As such, there is no outstanding issue of comparability. However, in attending to issues of comparability through the conceptualisation of a focus discourse community, a further important consideration emerges. Stated earlier in Section 2.3, a discourse community uses genres that are specific to that community. In the case of this study, the genre of the research article is the focus. However, it is not clear whether the economics academic research article in English, French, and Spanish constitutes one genre or three separate genres. If it is the former, then comparability across all three is not problematic; if it is the latter, then it is difficult to faithfully compare these genres and yield results of value if like with like is not being compared. In order to address this problem, the concept of genre and the research article as a cross-linguistic disciplinary genre are discussed in the following section.

2.4 Genre in academic writing

As with discipline, genre is an important variable in applied linguistic studies of academic discourse and, given the focus of this study on the discipline of economics in English, French, and Spanish, it is important to consider whether the genre attributed to that discipline in each language is comparable across these languages. Therefore, the goal of this section is to discuss the concept of genre and establish the research article as a

disciplinary genre that is comparable across languages. This is done through a theorisation of genre that helps identify the research article, the consideration of the research article in terms of both its textual properties and social and discursive context, and the operationalisation of the research article as a disciplinary genre. First, the concept of genre is discussed.

Understanding what constitutes a genre is integral to effective comparison in this study. This is because the fact that the texts in each language are referred to as research articles does not mean that they are inherently comparable. Rather, it is important to drill down and identify the features of the genre that will help ensure comparability across languages. However, this is no easy task, as genre has long been a contentious concept in the literature (Bawarshi and Reiff 2010, p.3) and the debate surrounding genre in language and rhetoric studies hosts three major schools of thought, namely: North American new rhetoric studies, Australian systemic functional linguistics, and English for specific purposes/languages for specific purposes. Looking to research centres beyond those focusing on English language genre studies, to these three one could also add the French and Swiss socio-discursive interactionism school of genre studies. Overall, all of these schools of genre study seem to differ largely ‘on the emphasis they give to context or text’ (Hyland 2002b, p.114), where text refers to the actual language in the document being studied such as the research article and context refers to the idea that all acts of writing are imbued with the social, professional, and cultural context of the writer (Hyland 1998, p.442). Next, the schools of genre are briefly compared in order to demonstrate the rationale for aligning this research with the language for specific purposes school.

New rhetoric is an American school that views genre in terms of social uses, where actions create a genre within situations and contexts (Miller 2003, p.21). Such a view analyses genre in educational contexts, where researchers such as Freedman (2013, p.124) have considered genre in terms of Bakhtin *et al.*'s (1986) dialogism and the creation of genre through social actions. This approach can help inform research on composition and literacy. Others have focused on rhetorical genre theory and the rhetorical situation (Paré 2014, p.86), which sees genre realised by an ‘exigence’ and constructed via ‘a text, an author, an audience, a purpose, and a setting or circumstance’. In the French and Swiss school of genre research, there is a notable overlap with new rhetoric, with similar

theorists such as ‘Bakhtin, Vygotsky, Wittgenstein, Foucault, and Habermas’ (Bawarshi and Reiff 2010, p.75) serving an important role in informing the socio-discursive interactionism therein (Bronckart 1997, pp.22-30). Socio-discursive interactionism, like new rhetoric, places an importance on the social and interactive side of genres (Baltar *et al.* 2009, pp.53-54) with much less consideration given to the text (Araújo 2010, p.46). Conversely, Halliday’s systemic functional linguistics and the Australian school place most importance on the texts being studied.

Halliday’s work in the Australian school has dealt largely with genre as a ‘staged goal-oriented, purposeful activity’ (Martin 2001, p.155) which views the context of situation (Hasan 2009, p.169), *i.e.* language use, and context of culture (Hasan 2009, p.169), *i.e.* the culture in which the language is used, as defining and inseparable characteristics. This reflects an important facet of this study, where language and culture appear inseparable as they influence the nature of the academic writing across languages. However, in systemic functional linguistics the focus is largely on the text and its structure and stages which characterise the text within the social context in which it used (Halliday 2009, p.59). This differs from this study, where importance is placed on the interdependence of text and context, where texts help create the wider discursive context and the wider discourse facilitates the creation of texts.

Overall, the new rhetoric and French and Swiss schools’ almost exclusive focus on context do not lend themselves well to a corpus analysis, where the text is as important as the context. Similarly, there are some issues in the application of systemic functional linguistics to this research since, firstly, systemic functional linguistics has typically focussed on primary and secondary school contexts (Martin 2009, pp.10-11); this is not reflective of the context studied here. Secondly, systemic functional linguistics places a much larger focus on functional rather than formal studies of genres (Webster 2009, p.1), and in this study the form is important for the cross-linguistic comparison. Moreover, the contrastive nature of this study brings with it a dynamic community that is culturally diverse, as evidenced in Section 2.2. This community takes the form of a focus disciplinary discourse community, as identified in Section 2.3., which is shaped by global trends and practices. Furthermore, while genres can help define that community, the community also defines their genres which are dynamic and reflexive, given the global

nature of the community. As such, the primacy given to text in systemic functional linguistics is incompatible with this study. Moving forward, the following discussion demonstrates that the English for specific purposes genre theories are most applicable to this research.

Although English for specific purposes genre theory in its formation began on studies of English (Hutchinson and Waters 1987, p.7), under the more multilingual-friendly term of language for specific purposes, important genre research has been conducted. This applies the English for specific purposes model to languages other than English (Martin 2010, p.3), such as French (Martin 2010), and Spanish (Sánchez-López 2012). Others have referred to this as foreign languages for specific purposes (Kordić and Cigan 2013). In this study, this model is henceforth considered the language for specific purposes model where, to-date, most research has been conducted on English or contrastively on English and other languages in academic writing contexts.

In the language for specific purposes school, genre places an importance on both context and text where the social interaction and the language used to achieve this are both balanced and important (Hyland 2002b, p.114), situating language for specific purposes genre studies in between those other schools mentioned earlier in this section. In this view of genre, texts usually have specific structures (Swales 1990, p.41), and research on texts often centres on their register and the role of specific linguistic feature like tenses, modality, and discourse markers within them. Context tells us more about the writers, their institutions, the social conventions surrounding their interactions with interlocutors, their cultures, and many more concepts that are embroidered into the fabric of their writing. However, while neatly presented here as two distinct concepts, in practical terms it is quite impossible to separate context from text, and text from context. The text cannot be produced without context, and the text cannot be understood without context. Likewise, context cannot be produced without specific texts that constitute it, *i.e.* discourse defines the context and the context defines the nature of the discourse (Bhatia 2004, p.24; Fairclough 2003, p.65; Hasan 2009, p.167). Research in this area has seen genres as flexible and dynamic (Bakhtin *et al.* 1986, p.105), where changes to genre are determined by changes perceived within texts and contexts throughout time and across linguistic and cultural spaces (cf. Section 2.2.1 of this chapter; see also Fløttum *et al.*

2006, pp.16-23, for a discussion of the role of culture on text and text on culture). In this way, genres are not fixed. Moreover, genres are seen to vary across disciplines (Bhatia 2004, p.23), which is unsurprising given the social nature of disciplinary discourse communities and their need to define the parameters of their memberships and communication, as discussed in Section 2.3.

In language for specific purposes, generic texts have ‘beginnings, middles and ends of various kinds’, which are analysed in terms of their ‘discourse structure’ (Swales 1990, p.41). Such analyses would require the interpretation of genres as ‘specifically discursive aspects of ways of acting and interacting in the course of social events’ (Fairclough 2003, p.65). More specifically, Bhatia defines genre in language for specific purposes as ‘essentially referring to language use in a conventionalised communicative setting in order to give expression to a specific set of communicative goals of a disciplinary or social institution’ (2004, p.24). Researchers in the language for specific purposes school have analysed genre both in terms of communicative purpose and social process (Bawarshi and Reiff 2010, pp.44-45). Communicative purpose refers to the purpose a genre fulfils (Bhatia 1993, p.13); the genre is constrained by this purpose, and if the purpose changes, it is likely the genre would too (Bhatia 1993, p.13). However, the role of communicative purpose has since been reconsidered, as communicative purpose is too complex to be awarded such a privilege and power as to constitute a genre (Askerhave and Swales 2001, p.209). For this reason, the concept of social process has become more important. In language for specific purposes, genre is constituted primarily by the social intention of the users and is later categorised by communicative purpose. In the view of language for specific purposes, genre has consistently been used to analyse academic texts, such as student and professional writing (Flowerdew 2005, p.323). This is often for the purpose of informing academic writing pedagogies and genre pedagogies in English as well as other languages like French and Spanish (Rodgers *et al.* 2011; Cesteros 2013; Bocanegra-Valle 2014; Tangpijaikul 2014).

Overall, this applied linguistic view of genre presents a general consensus that ‘genres are abstract, socially recognised ways of using language’ (Hyland 2002b, p.114) that provide writers with a framework to ‘achieve specific goals’ (Bhatia 2002a, p.16). To expand on this definition, genre is also viewed as dynamic and stable in the Bakhtinian

view (1986); genre, as a discourse object, is not static, but is in fact changeable and specialised in terms of language and discipline. Given the importance of genre to disciplinary discourse communities, as discussed in Section 2.3, the following section considers the research article in terms of its text and context in order to position it as a genre that is indigenised within individual disciplines like economics.

2.4.1 The research article as a disciplinary genre

What is clear from disciplinary studies is that discipline is consistently a cause of variation in academic writing – see for example Dahl (2004), where, regarding the use of metatext, discipline emerges as a greater variable than language. Therefore, it is important to consider whether the economics research article is a disciplinary genre in its own right, and one that spans the English, French, and Spanish focus disciplinary discourse community. Building on the language for specific purposes view of genre, this thesis follows Bhatia (2004, p.23) in recognising that a disciplinary genre is best understood as follows:

1. Genres are recognizable communicative events, characterized by a set of communicative purposes identified and mutually understood by members of the professional or academic community in which they regularly occur.
2. Genres are highly structured and conventionalized constructs, with constraints on allowable contributions not only in terms of the intentions one would like to give expression to and the shapes they often take, but also in terms of the lexico-grammatical resources one can employ to give discursive values to such formal features.
3. Established members of a particular professional community will have a much better knowledge and understanding of the use and exploitation of genres than those who are apprentices, new members or outsiders.
4. Although genres are viewed as conventionalized constructs, expert members of the disciplinary and professional communities often exploit generic resources to express not only “private” but also organizational intentions within the constructs of “socially recognized communicate purposes”.
5. Genres are reflections of disciplinary and organizational cultures, and in that sense, they focus on social actions embedded within disciplinary, professional and other institutional practices
6. All disciplinary and professional genres have integrity of their own, which is often identified with reference to a combination of textual, discursive and contextual factors.

Based on the above, the disciplinary research article can be seen as a communicative event, the purpose of which is to share, exchange, and discuss scholarly academic information and persuade readers of the value of research (Smith 2007, p.164). In the context of a discipline, this event is ‘mutually understood by members’ (Bhatia 2004,

p.23) of their discourse community. In the case of economics, it is reasonable to assume that the economics research article fulfils this criterion. This mutually understood communicative event belongs to the focus discourse community as, despite not sharing the language in which it is written, the research article genre originates in the epistemological empiricist paradigm as discussed in Section 2.2.1. Globally, this is more recognisable as a genre than outputs produced via the constructivist approaches to research dissemination typical of French and Spanish academic discourse.

In terms of the research article and its textual properties, the linguistic features of the research article constitute its register (Biber and Conrad 2009, p.7). Register is a contentious and polysemous term in the literature; for example, Halliday (1978, p.145); Yunick (1997, p.337); Biber and Conrad (2001, p.183); and Ramos (2015, p.65) offer differing perspectives on register. In this thesis, register is viewed as a variety of language that is categorised according to its usage (Syal and Jindal 2007, p.24), which varies according to situational context. Such variation is observed primarily in terms of lexicogrammatical features (Teich 2003, pp.23-24). Registers 'vary with the times, with fashion, and with ideological movements in society' (Lee 2001, p.47), and specialised registers like academic language tend to vary across languages and disciplines (Biber and Conrad 2001, p.183). Later in this chapter, Sections 2.5-2.7 illustrate that the research article behaves differently in different disciplines in terms of organisational patterns, use of metatext, and use of language across English, French, and Spanish. While the cross-linguistic differences may seem problematic, this is not necessarily the case, since, as previously established in Section 2.3.1, a focus discourse community is not limited by language. Furthermore, although the linguistic features may differ, the perception of the register of academic writing in each language sees academic writing as a formal written register (Swales and Feak 1994; Cassany 1999; Salager-Meyer and Ariza 2004) which renders the economics research article in English, French, and Spanish comparable. Moreover, the variation that does occur across discipline indicates that the economics research article, for example, is sufficiently different from other disciplines to be considered a disciplinary genre in its own right (Salas 2015).

From a contextual perspective, the disciplinary research article must respond to its own membership and employ acceptable means of interacting with the community when

compared to other disciplines. The discussion of the disciplinary discourse community in Section 2.3 and the interactive and social nature of research article discussed later in Section 2.8-2.10 support this view in the context of the economics research article. There is sufficient evidence that the economics research article contains idiosyncratic features that render it a reflection of the academic and disciplinary culture it represents; in other words, it is different from research articles in other disciplines (Fløttum *et al.* 2006; Lakić 2010; 2015; Salas 2015; Šćepanović 2015). Such features are discussed in Sections 2.8-2.10 and pertain to metadiscoursal differences where economics, for example, is seen to contain more metatext than linguistics or medicine (Fløttum *et al.* 2006, p.182). Seeing the reflexivity evident in the research article in its communication to disciplinary discourse communities both in terms of language choice and interpersonal practices, this thesis argues that the data studied herein contain a single disciplinary genre - the economics research article - which belongs to the economics focus disciplinary discourse community.

With the research article established as the genre on which the comparison of English, French, and Spanish academic writing is conducted, this research can be situated among studies of what has long been and continues to be ‘the main channel of [...] scholarly communication’ (Holmes 1997, p.322) in the academic community. The research article has received a lot of attention in research on academic discourse and, as discussed in Section 2.2, the research article is a genre that has spread to usage in languages other than English. That being said, there are a number of caveats. First, there is an evident dearth of research on academic writing in the economics academic research article. This is true for the English language and even more so for languages other than English where, in fact, there is much less research available that has studied research articles and academic writing in general in French and Spanish. The subsequent sections take a closer look at the textual features of the research article in English, French, and Spanish, drawing where possible from research on economics. However, given the sparsity of such research, research on other disciplines as well as academic writing in general has been included where necessary. In each case, the structure of research articles, the length of sentences and sentence types used therein, and the role of passivity, tense, and verbal modality in research articles are discussed, first in English, in Section 2.5, second in French, in Section 2.6, and third in Spanish, in Section 2.7.

Note that, as mentioned in Section 1.3, the choice of linguistic features listed above that form the basis of the description of the research article was based on an initial pilot analysis of the data. In the pilot study, these features appeared to impact the use of questions as rhetorical devices of reader engagement, and they appeared to differ across languages. Of course, there is much more that could be said about the research article; however, owing to limitations of space only these topics are considered in this thesis.

2.5 Academic writing and the English academic research article

Recognising the formal nature of the register of the English academic research article (Swales and Feak 1994), this section focuses on the textual elements of the research article genre. As such, this section presents an overview of research on the structure, sentence length, sentence type, the use of passivity, and the use of tenses and modal verbs in the English academic research article, in order to gain a clear view of its linguistic features. Where possible, attention has been given exclusively to the economics research article.

2.5.1 Structure

When considering the structure of the economics research article, the most typical format is described as IMRAD or IMRD which refer to the sections of Introduction, Method/Materials, Results (and Analysis), and Discussion (Swales 1990, p.134). Overall, research has shown that economics contains research articles that follow the traditional IMRAD format. However, this is not consistent, and studies have identified that, while this format is present, so too are other variations (Gómez *et al.* 1998, p.66; Fløttum *et al.* 2006, p.203). For example, different organisations of structures such as introduction, literature, methods, results and discussion, and conclusion, or introduction, methods, results and discussion, conclusion have also been found to occur in economics research articles (Lin and Evans 2012, pp.151-152). Fløttum *et al.* (2006, p.203) use the KIAP corpus to identify the structure of economics research articles as containing a title, introduction, thematic middle section, and conclusion, which reflect a more traditional humanities approach to research articles, and allow for consistency in the analysis of sections across the papers in each language. Given that this study uses the same KIAP corpus, the same view is adopted here for ease of cross-linguistic analysis. Overall, in terms of the overarching structure of the economics research article, studies show that its format tends to fall between what are generally referred to as hard and soft science

patterns *i.e.* to what degree they represent IMRD structures. However, this ambivalent position is potentially a result of the discipline of economics containing sub-disciplines that reflect the qualities of both the hard and soft sciences.

With regard to titles, economics and soft-sciences, it is argued, typically use two-unit titles to allow for a title that can both propose a ‘catchy’ first look and a detailed follow-up (Haggan 2004, p.297). In terms of titles, those in the economics research article tend to be long, and it has been observed that economics research article titles make use of non-finite verbs and colons to create titles containing pre- and post-modifiers (Gómez *et al.* 1998, p.72). Nagano (2015, p.137) describes titles in the economics research article in English as being shorter than other disciplines; they tend to use single-unit titles and comprise noun phrases and questions.

For introductions, economics uses longer introductions, which contain more steps and moves than traditional CARS⁵ model introductions. Of particular note is the need for a separate move that summarises previous research (Lakić 2010, p.91; 2015, p.46). Looking too at CARS models, Move 2 in economics research articles exhibits more examples of question raising than linguistics or medicine (Ryvitytė 2003, p.95). Furthermore, Dahl (2009, pp.384-385) finds that claims are not only signalled in results and discussion sections, but in the introduction and conclusion sections too. Carter-Thomas and Chambers (2012, p.32-33) identify a notable feature of introductions in economics research articles: in single-authored papers, first-person singular pronouns occur prominently. They find that 45.5% of all first-person roles are realised with “I”; notably, these are used to allow the writer to embody a researcher and writer role in the introduction.

Receiving less attention overall, results and discussion sections in the economics research article are shorter than similar sections in research articles from the fields of English language teaching, biology and computer engineering, for example (Khedri *et al.* 2013a, p.68). They also contain less interactional metadiscourse, but a higher frequency of attitude markers, when compared to those other disciplines. Like results and discussion sections, conclusions have received little attention. However, they do appear to be a space

⁵ The CARS (create a research space) model refers to Swales (2004, pp.203-232) conceptualisation of the moves of an introduction which serve to create a research space.

in which claims are signalled to the readers (Dahl 2009, p.385). Furthermore, Tutin (2010b, p.235) argues that conclusions in English academic research articles in economics make greater use of evaluative adjectives, such as those centering on importance, when compared to linguistics, for example. She also reports more examples of evaluative adjectives in introductions in economics than linguistics, but the difference is more significant in conclusion sections.

Overall, each section in the English economics research article appears to reflect idiosyncracies of the disciplinary discourse community. Next, research on the length of sentence and the types of sentences in academic writing and economics academic writing in English is presented.

2.5.2 Sentence length and sentence type

Academic writing in English is typically perceived as containing relatively short sentences. This perception is largely reflected in research. For example, English sentences are short in academic writing in abstracts (Van Bonn and Swales 2007, p.102). In fact, looking at economics research articles specifically, award winning writing in appears to exhibit shorter sentences, which has become a descriptive characteristic of good academic writing in English (Sawyer *et al.* 2008, p.114). This characteristic of shorter sentences in English is further bolstered by Ghivirigă (2012, p.4232), whose study of economics research papers written by first language English users and Romanian English users shows that the first language English users used shorter sentences than Romanian English users. This not only provides further evidence that English academic writing tends to have short sentences, but that sentence length is a potential challenge for those learning to write in the English academic community. This finding is reflected in Uzuner's work (2008, p.255), which argues that an issue with long sentences in English is representative of some of the problems encountered by those wanting to publish internationally in English. Looking at the Spanish users of English, Martín *et al.* (2014, p.65) undertook a survey of five Spanish universities' research centres, and claim that Spanish academics, notably from medicine backgrounds, feel they have problems with long sentences when writing for publication. As a solution, Martín *et al.* (2014, p.65) recommend a training course that explicitly deals with sentence length in English academic writing. Moreover, shorter sentence length has emerged elsewhere as a notable characteristic of English academic

writing in the social sciences, where ‘the type of revisions which were most often required by the journal editors [when reviewing papers written by non-first language users of English in the social sciences] had to do with the style and the discourse features of the texts, such as sentence length, complicated paragraphs, grammatical and lexical mistakes, and so on’ (Gea-Valor *et al.* 2014, p.53).

When considering sentence type, the contrastive analysis presented in this thesis centres on distinguishing between simple, complex, compound, and compound-complex sentences. The grammatical definition of these sentence types is presented in Section 4.3.2.8. Overall, there is little empirical research on sentence type in academic English. Of course, there are many style guides, such as Dunleavy (2003, pp.115-116), who advises for shorter sentences with simple and straightforward grammar. When using subordinating clauses, he argues for them to always be placed at the beginning or the end and should generally be avoided to reduce ambiguity. This will help keep sentence length below 40 words, and closer to the ideal 20 words, as he advises. Slightly more empirically informed style guides find that shorter sentences are typical of paragraph openers and closers (Northedge 2005, p.327). This reflects views of English academic writing as being culturally bound, more concise, and making use of shorter sentences (Bennett 2009, p.52). Along the same line of reasoning, Martín (2003, p.35) finds simple sentence types being used most noticeably in methodology sections in social sciences. That is not to say, however, that there is no place for complex, compound, and compound-complex sentences in English. In disciplines like medicine, more complex sentences are used in discussion sections (Li and Ge 2009, p.99), and Chen (2019, p.160) reports more compound sentences overall. Furthermore, there is evidence of reduced relative and full relative clauses occurring frequently in academic English in humanities and social sciences, which reflect more complex sentence types (Deveci and Nunn 2018, p.22).

Overall, the research on academic writing in English offers a general consensus that English academic writing generally favours simpler sentences. Furthermore, the clear message from research on academic writing in English, and specifically in economics, is the use of short, simple, and clear sentences which has been evidenced in the discussion in this section. The following section considers the role of passivity in academic writing in English.

2.5.3 Passivity

Passivity and voice have long been the studied within research on academic writing in English. Studies have found that active voice coupled with “we” allows for the writer to discuss their current work, and active voices with others is used to mention non-conflicting work (Tarone *et al.* 1981, p.125). Interestingly, in Tarone *et al.*’s work, the passive is used to discuss conflicting work as well as future work and also has an emphatic function (1981, p.125). The literature also suggests that tense impacts the use of passive; present perfect passives tend to refer to someone else’s work and past simple passive to the writer’s own work (Shaw 1992, p.311). Moreover, the choice of passive, it is argued, is often related not only to agency but to the thematic position of the information. That is to say, word order can be modified to achieve greater cohesion (Shaw 1992, p.311). Later, Flowerdew (2012, p.3) identifies this same occurrence and argues that sentence construction may dictate the need for passives. This is long recognised in the literature, where ‘the higher a constituent ranks on the two hierarchies the greater the likelihood of functioning as the topic (*i.e.* the initial sentential element) of the clause’ (Siewierska 1986, p.221). Lakić’s (1997, p.8) study, which uses Swales’ CARS model to analyse introduction sections of economics research articles, finds that the active voice is most used, and, again, that there is a relationship between tense and voice. Out of 200 verbs studied, 10.5% were present simple passives, 2.5% were past simple passives, and 12.5% were present perfect passives. In some cases, he finds that these passives function to describe the present research *i.e.* ‘In this paper a model... is proposed’ or to ‘stress the importance of the research itself, not their contribution to it’ (Lakić 1997, p.12). He also finds evidence of passive being used to announce findings. Martínez (2001) presents an interesting study of English experimental papers, in which she argues that passive voice is used to create impersonality and a distance between the writer and the information they present. In her study, she finds that passive voice is most frequently used in methods sections of research papers in physical, biological, and social sciences. In the social sciences, which are most relevant here, she finds that 54% of the voice in methods sections are passive. This compared notably to 10% in introductions, 23% in results, and 15% in discussions. This usage of passives has been reported elsewhere, where passivity in methodology sections is common in English (Martín 2003, p.36). However, this use of passivity is explicable, not only in terms of sentence construction, but in terms of rhetoric

and register. The use of passivity, which has also been identified as an integral feature of research article abstracts in English, ‘can be attributed to [...] the English espousal of “I” [and] cultural expectations as to what constitutes appropriate formal academic style, particularly with regard to how an author is expected to refer to herself or himself’ (Van Bonn and Swales 2007, pp.104-105). As in the case of others mentioned earlier, Guerra and Guerra (2008, p.97) find a correspondence between passive voice and first person pronouns, and Bondi (2010, p.226) finds too that, in economics research articles, there is a tendency to use passive voice to introduce claims. From an interpersonal perspective, passive voice is also known to have a function for creating passive stance (Baratta 2009, p.1414). More recently, Amnuai (2019, p.5) finds that the use of passive voice in accountancy papers was very infrequent when compared to other disciplines and again occurred in methods sections primarily.

Overall, despite researchers like Day and Sakaduski (2011, p.125) arguing for the avoidance of the passive voice, it has become clear that, although infrequent, the passive voice does hold value for many reasons in academic writing. Moreover, in the context of economics, the passive voice is largely constrained to methodology sections. The following section presents research surrounding tense and verbal modality use in academic writing in English.

2.5.4 Tense and verbal modality

English academic writing in research articles has been categorised as being ‘written with full sentences that always have verbs, [the tenses for which] are determined by the type of context in which they are used’ (Deny *et al.* 2017, p.229). These tenses are challenging for writers (Wallwork 2011, p.132), not just owing to form, but also because their relationship to function and their misuse can confuse readers. For example, when considering tenses in academic writing, past tenses have been seen to predict a lack of generality, while present perfect and present simple tenses are used to make more general claims (Lackstrom *et al.* 1972, p.5). Building on this work, Oster (1981, p.89) sees present perfect as a resource with which writers can build on previous research. He sees past simple as being used to discuss contentious literature with which authors disagree, and he finds that present simple is used to create positive relationships between existing research and proposed research findings, or express research in a more neutral manner. This

research is corroborated by others such as Malcolm (1987, p.37), who finds that the present tense is useful for making generic claims while, interestingly, the past tense refers to specific research, and the perfect aspect allows writers to address general areas of research.

Looking deeper into the complexities and nuance of tense use in academic writing, past tenses have been found to be used to create supportive relationships between ideas (Shaw 1992, p.316) and, in the introduction sections of economics research articles, tense distribution has been documented by Lakić (1997). Lakić (1997, p.9) finds that 55% of verbs took present simple tenses, 14.5% used past simple, and 30.5% used the present perfect aspect. His work showed that present simple is used to refer to research that supports the writers' own work. He also finds that 'the authors of economics research articles do not only accept results of other authors but put forth their own models on which economics as a field of science is based'; and to do so, they make use of the past simple. Moreover, he sees the use of the present simple and the present perfect as being common among the presentation of facts in economics, where existing information is 'generally accepted and new knowledge is built upon' using the present perfect (Lakić 1997, p.9).

The role of time is complex in academic writing; while it may be reasonable to assume that the past simple refers to something in the past, tense is not always used to represent temporal distance, as demonstrated by the abovementioned studies. Further studies show that the present tense can be used to refer to past research (Swales and Feak 2012, p.345) and overall, past tenses are generally less common in academic writing (Biber *et al.* 1999, p.456), especially in the humanities (Winkler and Metherell, 2012, pp.108-109). This may be surprising, given that all references to literature in research articles must refer to past research that is both established and published. Therefore, tense does not necessarily correspond to time, but rather, as this discussion has illustrated, to writers' perceived relationships with others' work. Importantly, however, this does vary according to discipline (Deny *et al.* 2017, p.233). For example, Deny *et al.* (2017, p.233) show that academic research articles in the social sciences make greater use of the present simple tense than any other disciplines. In fact, 73.6% of all verbs are in the present tense. This contrasts strongly with the health and life sciences, where only 38.5% and 55% of all verbs are in the present tense. Furthermore, this behaves similarly to the physical sciences,

where 75.3% of all verbs are in the present. For past simple, in the social sciences only 22.4% of all verbs in past simple tense. This is similar to physical sciences, where 19.2% of verbs are past simple, and differs greatly from health and life sciences, where 57.8% and 41.3% of all verbs are in the past simple. Moreover, Deny *et al.* find evidence to suggest that location predicts different verb tenses; in social sciences, the present simple occurs in all sections, but especially in the background and the conclusion sections, while the past simple is more frequent in results sections.

As is clear from this summary, much of this research has focussed on present simple, past simple, and present perfect. In lieu of explicit tense, modality in English has been found to offer subjective views and perspectives (Parkinson and Adendorff, 2005, p.288), and relays evaluations conceptualised in possibility, probability, permission, volition, obligation, and necessity (Biber *et al.*, 1999, pp.485-486). As such, modal verbs can function similarly to tenses, whereby certain modal verbs can signal disagreement with previous research by expressing doubt *i.e.* “might” or “could”. Effectively, this means that modality allows writers to encode extra information, such as their opinion, in their writing (Carrió-Pastor 2014, p.155). While it is beyond the scope of this thesis to engage with the theoretical foundation and complexity of the field of modality, what is interesting is the relevance of modality to academic writing in the academic research article. For example, Vázquez and Giner’s (2008, p.180) study of epistemic modality⁶ markers as hedges in marketing, biology, and mechanical engineering research articles finds that such hedges are twice as common in marketing as biology and three times as common in marketing as mechanical engineering. However, of course, their study captures more than just modal verbs.

Looking at the role of modal verbs specifically, English has been found to use many modal verbs, such as “can”, quite frequently in academic writing to show possibility (Kanoksilapatham 2012, p.16), and business case studies use modal verbs only second to the present simple (Nathan 2013, p.64). Modals such as “would” are particularly salient (Nathan 2013, p.64) and are used to propose hypothetical future actions and consequences. Similarly, modals like “may” are found to be most used in introduction and discussion sections of research articles in economics in order to hedge claims and

⁶ Epistemic modality refers to a subfield of modality that captures evaluative modality.

contributions (Deny *et al.* 2017, pp.241-242). Overall, the role of modal verbs in academic writing, and their link to interpersonal metadiscourse, stance, and hedging, positions modal verbs as a valuable resource for English academic writing – for more detail on topics like metadiscourse see Section 2.8 and for a discussion of stance, see Section 2.9

2.6 Academic writing and the French academic research article

Building on the established knowledge of the English language academic research articles in the previous section and recognising the formal nature of the register of academic writing and the academic research article in French (Salager-Meyer and Ariza 2004), this section presents an overview of research on the structure, sentence length, sentence type, the use of passivity, and the use of tenses and modal verbs in the French academic research article, in order to gain a clear view of its textual and linguistic features.

It is important to note that, while the scarcity of research specifically on the economics research article in English is noted in Section 2.5, this scarcity expands to the research article in general in the context of French and Spanish. Because of this, this section is somewhat shorter and less specific to economics than the section on research on English presented in Section 2.5. Moreover, this section and its subsections draw on research on academic writing in French more generally and are not limited to the research article. Nonetheless, this section presents relevant research on the French academic research article and academic writing in French that remains valuable to this study.

2.6.1 Structure

When considering the structure of the economics research article, Liddicoat (1992; 1993; 1995; 1997) has made important contributions to the description of the French academic research article. He argues that French, like English, exhibits the typical IMR(A)D format and that the movement of information within the research article is linear, not digressive (1997, p.15). However, subsequent studies have shown that French exhibits inconsistencies in this format in the economics research article, whereby not all texts contain explicit conclusion sections, for example, and do not always follow the IMR(A)D schema (Fløttum *et al.* 2006, p.203). Regarding titles in research articles, concise and precise titles, which capture the essence of the paper, are encouraged in literature that offers guidance to writers and practitioners (Lettrilliart and Schott 2007, p.631).

However, more descriptive studies have found that titles in French research articles can be problematic. This is because the way they are written can make them ambiguous or unclear, which leads them to being poorly categorised in large international bibliographic databases (Alès *et al.* 2016, p.29).

Features of particular sections have also been noted by Carter-Thomas and Chambers (2012, p.27), who find that French uses very few first-person pronouns in the introduction sections. Furthermore, Tutin (2010b, p.235) finds that conclusions in French economics research articles make little use of evaluative adjectives. Others have found that methodology sections are the only sections to make any notable use of imperfect past tenses (Liddicoat 1997, p.16), and economics bibliographies in French are particularly short, as there is not a major need for substantiating claims made by referring to a selection of research papers (Grossmann and Tutin 2010b, p.294). Overall, there is little research on the generic structure of the economics research article in French, but research on research articles in general indicates some use of IMR(A)D formats. Evidence of inconsistencies in the genre with divergence from IMR(A)D formats, as well as evidence of the use or non-use of particular linguistic features, such as evaluative adjectives, pertaining to specific sections of the research article, such as the conclusion. The next section considers sentence length and type.

2.6.2 Sentence length and sentence type

With regard to sentence length in academic French, guidance for writing within the French academy suggests that texts should be neatly written in French, with clear yet elegant wording, and should make use of concise and short sentences with a maximum length of 16 words (Benichoux *et al.* 1985, p.53). Empirical studies of written academic French support this view; for example, in a corpus study of 150 texts in 6 scientific disciplines, the average sentence length was 15.97 words (Liddicoat 1997, p.16). However, the claim that French academic writing should be neat and clear has been rejected elsewhere. This is because clear word order, cohesion, and coherence have been argued to be uncharacteristic of French writing (Birch-Bécaas 1996, p.77), which has been found to also include very long sentences (Van Bonn and Swales 2007, p.102).

In terms of sentence type, Benichoux *et al.* (1985, p.53) advise the use of the subject-verb-complement sentence structure when writing in academic French, and for clarity,

simplicity, and readability in writing, each sentence should begin with the most important fact. However, verbless sentences are a documented feature of academic writing in romance languages like French (Bennett and Muresan 2016, p.100). As opposed to documented simplicity and readability, romance languages like French are seen in the literature as complex and literary in their sentence style, and so they are said to make use of long and non-linear processes. Such languages, it is argued, are notorious for the complexity of their writing style; Bennett and Muresan (2016) claim that ‘French [has] been described as so preoccupied with linguistic artistry that they are prepared to sacrifice scientific rigour to achieve balance and symmetry in their prose’ (p.96). Moreover, in terms of sentence construction, French makes use of various combinations of coordinating and subordinating clauses – see grammars such as Collins (2011) and Hawkins and Towell (2015). However, unlike English, academic French appears to have a greater penchant for subordination (Van Bonn and Swales 2007, p.99); as such, descriptive studies of French academic writing tend to identify French sentences as long, heavily subordinated, and complicated. Next, passivity, tense, and verbal modality in French academic writing are presented.

2.6.3 Passivity, tense, and verbal modality

In terms of passivity, while the passive is infrequent in general in French, it appears to be unusually frequent in academic written French, where it is used to place certain information, like agency, in the background (Liddicoat 1997, p.15). However, in research article abstracts, there is no strong evidence of passive use in French academic writing (Van Bonn and Swales 2007, p.100), confining its presence to the research article proper. These studies focus specifically on passive voice, and in a wider view, passivity can also be evoked through the use of the ‘*on*’ pronoun or the ‘*se*’ reflexive, which allow the writer to behave impersonally and hide the agent of the text (Collins 2011, p.122; Fløttum *et al.* 2006, p.115). There is a sparsity of research on passivity in French, and the research that does exist is ambivalent on its role and purpose in French, not least in the context of economics academic writing. Overall, despite this sparsity of research, passivity appears to be a feature of French academic writing.

In terms of tense in French academic writing, Benichoux *et al.* (1985, p.54) advise, prescriptively, the use of the present tense to present information and the use of the

imperfect to present contributions. They also suggest the use of present or the past tense, *passé composé*, to cite others, and the imperative to make recommendations. However, the empirical landscape in the literature appears more complex; tense in academic writing does not necessarily have any relationship to time, and like in English in Section 2.5.4, can serve a range of rhetorical features (Liddicoat 1997, p.15). The present and the *passé composé* have been found to be the most frequently used tenses; the present is used to refer to general information and the *passé composé* is used to refer to existing research (Liddicoat 1997, p.15). Furthermore, past tenses like the imperfect are infrequent in French academic writing and the imperfect is usually confined to methods sections (Liddicoat 1997, p.16). There also appear to be further functional values of tense in French academic writing, where the *passé composé* is used to present new methods; conditionals present new findings and show evaluation; and present tense is used to present a new result (Liddicoat 1997, p.16). Moreover, the present and *passé composé* are both used to cite others' work. In such instances the present typically refers to studies considered more important, while the *passé composé* is used to refer to more peripheral studies. However, Liddicoat finds little to no evidence of imperatives being used in French academic writing (1997, p.16). Interestingly, in their work on research article abstracts, Van Bonn and Swales (2007) report no distinguishing feature of tenses in French research article abstracts, while Burrough-Boenisch (2003, pp.7-8) identifies that the historic present is a very common feature of academic French. This historic present tense is used to refer to the past. For the perfective aspects, present perfect forms in French have been found to be unusually frequent in academic writing when compared to less specialised discourses (Grossmann and Tutin 2010b, p.287).

Regarding verbal modality, Benichoux *et al.* (1985, p.55) recommend that academic writers avoid imprecision, which Liddicoat (1997, p.16) argues is a recommendation for the avoidance of hedging. In terms of hedging, modal verbs have been found to be important in French academic writing (Liddicoat 2005, pp.13-14), which questions Benichoux *et al.*'s recommendation (1985, p.55). Becker and Remberger's (2010) edited volume on modality and mood recognises throughout that the subjunctive can act as a form evaluative modality. As such, not only modals but verbs in the subjunctive mood are used to hedge and create a form of verbal modality (Liddicoat 2005, p.13). This means that subjunctive moods, as well as recognised modal verbs like *sembler*, *paraître*, and

pourrait, are common means of expressing verbal modality in French academic writing (Vold 2006, p.69); however, it is important to note that modality is less common in French overall, especially in the soft-sciences (Vold 2006, p.69). Overall, passivity and modality appear to be features of academic writing in French and tenses specifically offer rich variation in use beyond temporal implications.

2.7 Academic writing and the Spanish academic research article

The third and final language under consideration is Spanish. Recognising the formal nature of the register of Spanish academic writing and the academic research article (Cassany 1999), this section presents an overview of research on structure, sentence length, sentence type, passivity, tense, and modal verbs in the Spanish academic research article, in order to gain a clear view of its linguistic features.

As with French, there is much less research on the nature of Spanish academic writing in economics, and the academic research article in general. In fact, Oliver del Olmo (2015) recognises that only very recently has attention turned to studying academic writing in Spanish, owing to increased interest in interlinguistic and intercultural use of rhetoric. Therefore, as with French in Section 2.6, in addition to studies on the Spanish academic research article, general views of Spanish academic writing are presented, where necessary, to generate a more comprehensive and relevant review.

2.7.1 Structure, sentence length, and sentence type

When considering the structure of the academic research article in Spanish, it has been argued that Spanish is not very linear in the organisation of information within the text of the research article, which does not always follow a consistent path (Mur-Dueñas 2011, p.3072). However, like French in Section 2.8.1, although inconsistencies do occur within the generic structure of the research article, Spanish research articles do appear to reflect an IMR(A)D approach (Lorés-Sanz 2011a, p.178), at least in research articles on business management. However, there remains evidence of idiosyncratic behaviour, where exclusive pronouns are most common in methodology sections in business management articles (Lorés-Sanz 2011a, p.178), for example. Other researchers identify the role of passives and the past tense as being particularly common in methodology sections (Martín 2003, p.36; Lorés-Sanz 2011b, p.10) and find that passives feature prominently in discussion sections (Pérez-Llantada 2010, p.64). There is also evidence of less

participant-oriented metadiscourse in introductions than discussions (Pérez-Llantada 2010, p.47). In terms of titles, Soler (2009, p.54; 2011, pp.130-131) identifies that titles in Spanish research articles tend not to be full sentences, but make use of compound titles, and tend to be long. Overall, while the IMR(A)D format is present in Spanish research articles, Breeze (2012, p.28) has found that academic genres in Spanish tend not to be so fixed and tend to be more open to variation. This overall discussion has shown that internally, research articles in Spanish support a range of formats.

In terms of sentence length and type, Spanish academic writing has been seen to exhibit very long sentences (Breeze 2012, p.27). Similarly, LoCastro (2008, p.205) and Perez and Ramiro (2015, p.578) find long sentences to be characteristic of Spanish writing. However, it is of note that others have found that the average Spanish sentences are not any longer than English in academic contexts (Taft 2011, p.512). This similarity with English is explicable, as Spanish academic writing has been seen to behave increasingly like English (Salager-Meyer *et al.* 2003, pp.239-240). In terms of sentence type, like English, Spanish grammars identify the evidence of sentence types like simple, complex, compound, and compound-complex (Pountain and Kattán-Ibarra 2004), and research has exhibited evidence that complex sentences and subordination are particularly characteristic of academic Spanish in the context of business disciplines (Pérez-Llantada 2010, p.55). Overall, this is unsurprising given that, as with French, Spanish has been labelled complex and literary in terms of sentence style with evidence of long and non-linear processes (Bennett and Muresan 2016, p.96). Overall, Bennett and Muresan (2016, p.95) consolidate research on Spanish academic writing by saying that ‘Spanish academic prose has been described by linguists and educators as “elaborate”, “ornate”, “repetitious”, with long “flowery” sentences and a penchant for subordination’.

Therefore, like French, the genre of the Spanish research article is more likely to vary in format, not only in terms of organisation but also in terms of sentence length. Moreover, it most likely contains complex sentences, with many subordinate and non-linear clausal components. The next section discusses the role of passivity, tense, and verbal modality in Spanish academic writing.

2.7.2 Passivity, tense, and verbal modality

In terms of passivity, two types of passive emerge in Spanish academic writing. The first passive is constructed with a form of the verb *ser* and the past participle. The second is constructed with *se* followed by a verb in the active voice (Espinoza 1997, p.231). Interestingly, the *se*-passive appears to be more common in academic writing and usually occurs in the third person single or plural (Espinoza 1997, p.231). Passivity, while under-researched in academic writing in the economics research article, has been seen to have specific functions in the likes of biomedical research articles, where it is used to introduce topics and remind readers of previous material (Pérez-Llantada 2010, p.53). Others have observed the role of passives in reporting methodology (Martín 2003, p.36), and Oliver del Olmo (2005, p.213) finds that the passive is used to hedge in medical discourse. More recently, the use of the *se*-passive in Spanish medical articles has been seen as an important rhetorical device for demonstrating impersonality (Salazar *et al.* 2013, p.129).

Similarly, the role of tenses in the economics research article is somewhat unclear, although generally, in academic writing, there is a preference for simple tenses and perfect tenses in particular (Espinoza 1997, pp.239-242). This is similar to French, as discussed in Section 2.6.3. Spanish writers appear to use the past tense specifically in methods sections with a predominance of present tenses overall (Martín 2003, p.36). Furthermore, looking at methodology sections in more detail, there is evidence of past, present, present perfect, and future tenses being used for various reasons in business management research articles (Lorés-Sanz 2011b, p.10). Elaborating on these functions, Lorés-Sanz finds that the past is used for past acts, the present is used to describe processes, the perfect is used for a process in the past that relates to the present, and the future is used to plan. While limited research exists on Spanish academic writing and tense, tense appears to perform functions beyond references to time, as in French and English in Section 2.5.4 and Section 2.6.3 respectively.

In terms of verbal modality in Spanish, like French, the subjunctive mood behaves like modal verbs in English (Collentine 2014, p.272), and such modality has been found to create impersonality (Pérez-Llantada 2010, p.55) or mitigate face threatening acts (Lorés-Sanz 2011a, p.189). Lorés-Sanz (2011b, p.12) also finds that modal verbs in Spanish are used to express stance in terms of certainty, obligation, and permission. Overall, while

there is a lack of research in this area, especially with regards to the economics research article, what emerges from the research discussed is that, like French, Spanish makes particular use of passives, tenses, and verbal modality to achieve functional ends in academic writing.

Section 2.5-Section 2.7 have discussed the research article and academic writing in English, French, and Spanish. Overall, this review of the literature serves to illustrate the linguistic features of academic writing in research articles in each language. From these discussions, it is clear that the internal structure of research articles is flexible and changeable, and the academic writing of the economics research article differs from other disciplines in each language. This reflects the importance placed on dynamism in defining genre and the operationalisation of the economics research article, as discussed in Section 2.4. The varied sentence length and sentence types that are found in research articles in each language also reflect the cultural context discussed in Section 2.2.2, where there are implications on the concept of reader-responsible and writer-responsible languages. Furthermore, the role of passivity, tense, and modality in serving idiosyncratic functions in the academic research article demonstrate important textual differences that are only understandable within their academic community. Overall, as mentioned in Section 2.4, the text and the context are inseparable, and the many textual features discussed appear to have contextual implications. Looking towards the context, the following section moves beyond the individual linguistic features of academic writing and the academic research article in English, French, and Spanish, and instead moves to elaborate on the interactive nature of the research article by considering its metadiscoursal features in terms of evaluation, stance, and reader engagement.

2.8 Metadiscourse in academic writing and the academic research article

While the elements discussed in the previous section describe the textual features of academic writing and the academic research article in English, French, and Spanish, one key facet of this description is missing, namely the role of the genre of the research article in contributing to the wider context as a social and interactive discourse. To this end, this section considers the interactive nature of the research article, outlining its interactive and interpersonal roles in order to demonstrate the importance of the research article, beyond the textual level, to its academic discourse community. In order to do this, first, theories

of interpersonal communication are discussed. Second, of these theories, the concept of metadiscourse is adopted to show how research articles perform their interpersonal activity. Following that, research on metadiscourse in the academic research article in English, French, and Spanish is discussed in order to illustrate the current state of the art. Overall, the goal of this section is to substantiate the view that the research article is an interpersonal and social text. This is important when considering a core facet of this thesis centres on reader engagement, a concept which presupposes interaction – for more on the concept of reader engagement see its detailed discussion in Section 2.9.1.

Within the field of academic discourse, it is widely accepted that ‘implicit in every act of academic communication is the writer’s awareness of the social context and professional consequences of the writing’ (Hyland 1998, p.442). This too pertains to the genre studied herein, the research article, which in its very nature is rhetorical, as it incorporates both writers and readers within its text (Hunston and Thompson 2000; Mauranen and Bondi 2003) and also ‘presupposes the active role of [its] readers’ (Fløttum *et al.* 2006, p.182). Broadly speaking, researchers investigating interaction in research articles have benefitted from studying language in terms of various theories of interpersonal communication. In this section, the core theories in the field are discussed, with a particular focus on their role in the study of academic writing in English, French, and Spanish.

Among theories of interpersonal communication in writing, the enunciative approach has gained distinction within research on French and Spanish. Enunciation considers the way in which subjectivity is instilled in writing through the ‘enactment of language by an individual act of utilization’ (Benveniste 1974, p.80). Many studies incorporating the enunciative approach to interpersonal communication conduct polyphonic investigations of ‘voices’ in written texts, where multiple voices are enacted (Angermüller 2011). Polyphony as a concept has undergone many theorisations which are heavily debated (Dendale 2006, p.3). However the concept of polyphony generally views a text as a ‘scene or a drama of interrelated voices or points of view attributed to more or less abstract instances in addition to the speaker or writer’ (Fløttum *et al.* 2006, p.35); within this approach, researchers such as Ducrot (1984); Nølke *et al.* (2004); Fløttum *et al.* (2006); Fløttum *et al.* (2013) have undertaken research on this concept of voice, seeing ‘nested

voices' as 'chained together in light of their argumentative value' (Amossy 2005; Angermüller 2011, p.2994). In this way, writers, through enunciation, can inform their readers as to who is speaking (Rabatel 1998, p.125).

Conversely, although theoretically linked, other studies reject polyphony and seek answers within the French *praxématique* school (Bres 2001), which argues that polyphony is an ill-named concept that was initially entitled dialogism by Bakhtin (1986) (Dendale 2006, p.21). These schools differ in their juxtaposition of voices; the *praxématique* school views dialogue as a free interaction (Bres 2001, p.22), as opposed to a hierarchical polyphony as in the enunciative approach. However, these schools of thought would not necessarily allow this thesis to answer the questions it seeks to answer. The enunciative approach places importance on the many actors and voices present, and considers the other players, voices, actors and institutions evoked by the writers, beyond the writer and the reader, as identities in their own right (Angermüller 2011, p.2993). In this way, the enunciative approach moves beyond the social role of genre within a disciplinary discourse community, to the placement of this community within its socio-cultural context (Angermüller 2011, p.2993), and therefore is not applicable in this study, given its focus study on disciplinary discourse.

A further theory of interpersonal communication that has gained prominence in the field is appraisal theory, which is situated within the Australian school of systemic functional linguistics (Martin and White 2005, p.33). This theory places semantic analyses and functions at its core (Webster 2009, p.1) and encompasses varying foci on "attitude", "engagement" and "graduation". Attitude illustrates opinions, engagement allows for the presence of voices beyond that of the author, and graduation seeks to modify the strength of attitudinal and engaging efforts (Read and Carroll 2012, p.423). In this way, appraisal theory allows for the linguistic analysis of writer stance and reader presence (Martin and White 2005, p.152) in ways that differ to metadiscourse. However, an important difference between metadiscourse and appraisal theory is that, in systemic functional linguistics, propositional, interpersonal, and textual elements of texts must each be considered together and holistically (Schulze 2009, p.82), whereas in metadiscourse, they can be taken independently (Hyland 2005b, p.27). Moreover, appraisal theory has traditionally focused on texts with larger audiences and participants (Martin and Rose

2007, p.93), such as journalistic texts (McCabe and Heilman 2007) and political speeches (Guan *et al.* 2006). Hyland argues that, because appraisal theory has been applied to mass-audience texts, it does not serve as a ‘model of evaluative discourse’ for academic writing, as it has derived from the typical features of academic writing (2005a, p.175).

Overall, though disparate in nature, all of these approaches share an overarching theme in interpersonal interaction and have been applied to research on English (Santos and da Silva 2016), French (Donahue 2008), and Spanish (García Negroni 2008) academic writing. Another important theory of interpersonal communication, and one which has greatly informed this study’s understanding of interpersonal communication in academic writing, is metadiscourse.

Metadiscourse is a view of interpersonal communication, which attends, to some degree, to each of the schools of thought identified above. Although an ambiguous term and has been referred to as ‘discourse about discourse’ and ‘talk about talk’ (Hyland 2005b, p.16), it is, in its essence, concerned with the ways in which writers impose themselves on a text with the express intention of manipulating or anticipating the readers’ inferences of the text (Hyland 2005b, p.18). Metadiscourse is a framework for analysing the ways in which ‘communication as social engagement’ (Hyland 2010, p.127) is achieved through writers’ positioning of themselves and their readers (Hyland 2010, p.127). Hyland’s framework considers all elements of metadiscourse as interpersonal (Li and Wharton 2012, p.347), unlike systemic functional linguistics, for example, which does not consider the textual metafunction as interpersonal. Moreover, beyond research on English, metadiscourse has also been employed in studies of languages other than English, for example Mur-Dueñas (2011) in Spanish and Hempel and Degand (2006) in French. This is an important consideration, as this study also applies to English, French, and Spanish academic writing.

Broadly speaking, metadiscourse sees academic genres such as the research article as more than just the delivery of propositional content. Rather, it sees the role of the article as embedding such content with ‘personalities, attitudes and assumptions of those who are communicating’ (Hyland 2005b, p.3), for the express purpose of influencing the readers. This influence usually conveys, in part, the author’s person and credible nature as researcher and expert in terms of their relationship to the message (Hyland 2004b, p.20); a message which, seen through the lens of metadiscourse, is never truly neutral and

is always value-laden (Hyland 2005a, p.185; Hyland 2005b, p.4). Moreover, Hyland has been found to argue that metadiscourse allows for a more ‘reader-friendly’ text (Amiryousefi and Rasekh 2010, p.164). However, as discussed in Section 2.2, such a description must be understood as culturally relative. In Hyland’s case, he was discussing English academic writing. Nonetheless, seeing metadiscourse as a means to perceive the wider social context in writing allows us to better understand not only the register and textual features of a text, such as the research article, but to also better understand the social practices, conventions, and identities of writers within a disciplinary genre and their discourse community. Metadiscourse is concerned with non-propositional content, both discourse organising and reader-writer interaction, and can be subcategorised as textual or interpersonal (Hyland 2005b, p.26).

Textual metadiscourse allows writers to assist readers to ‘connect, organise, interpret, evaluate, and develop attitudes towards’ (Vande Kopple 2002, p.93) propositional content. This is done through the use of ‘textual markers’ (Hyland and Tse 2004, p.163) that create connections, sequences, reminders, and announcements, for example (Hyland 2005b p.32). Importantly, such markers are simultaneously textual and interpersonal; textual markers serve to anticipate reader needs and direct the reader persuasively (2005a, p.185; Hyland 2005b, p.27). Beyond the textual metadiscourse, interpersonal metadiscourse is less concerned with the organisation of the text, and more with writers interacting directly with their readers through the use of hedging and boosting (Orta 2010), attitude and certainty (Rentel 2006), questioning (Curry and Chambers 2017), and directing (Hyland 2004b), for example. That being said, such interpersonal practices can also have text organising effects. This study focuses on interpersonal metadiscourse and develops the body of work on evaluative metadiscourse, which is discussed in Section 2.9.

In terms of studies on textual and interpersonal metadiscourse in English, French, and Spanish, there exists an important body of work. For example, in English, economics research articles have been found to be more at ease with expressing doubt and uncertainty than disciplines in the hard sciences (Gómez *et al.* 1998, p.72). More specific research on interactive linguistic features of writing has also proposed that transition markers such as “furthermore”, “equally”, “in the same way”, “thus”, “therefore”,

“however”, and “in contrast” occur more frequently in economics research articles than applied linguistics articles, for example (Khedri *et al.* 2013b, p.322). Dahl (2004, p.1818) investigated writer presence in academic research articles and found that metatext occurs frequently in English economics writing, and much more so than in French. This is unsurprising given that French has long been identified as impersonal; Loffler-Laurian (1980, p.135), for example, describe academic writing in French as a discourse which uses all resources available to maintain impersonality:

D’une manière générale, le français évite absolument la première personne du singulier, emploie le moins possible la première personne du pluriel, et se trouve amené ainsi à utiliser toutes sortes de formulations dites impersonnelles, réfléchies ou passives

Furthermore, French has been found to show a more critical stance, which is evidence of the importance of metadiscourse in French (Van Bonn and Swales 2007, p.104). Others have argued that French is not very explicit (Fløttum *et al.* 2006, p.184) but, in what seems like a contradiction, is very authoritative (Salager-Meyer 2011a, p.36). Studies on business management research articles shows that Spanish uses fewer evaluative connectors (Mur-Dueñas 2007a, p.43) and, in a later study, Mur-Dueñas (2011, p.3071) finds that metadiscourse in English business management research articles occurs so frequently that interpersonal elements of academic writing, she argues, are characteristic of the discipline. In contrast, Spanish, seems to make less use of exclusive pronouns (Lorés-Sanz 2011a, p.179) to exhibit writer identity, and it has been argued that metatext is infrequent in Spanish generally (Valero-Garcés 1996, p.291). Where metatext does occur, it has been found that Spanish contains fuzzier cohesion devices and more vague connections in academic writing such as the over use of demonstrative pronouns (Breeze 2012, p.28). This view also reflects the position of Spanish as being more reader-responsible than English. From an interdisciplinary perspective, Salas’ (2015 p.34) study of metadiscourse in Spanish research articles in economics, medicine, and linguistics finds that economics and medicine contain significantly fewer metadiscourse markers than linguistics.

In summary, metadiscourse ‘refer[s] to aspects of a text which explicitly organise the discourse, engage the audience and signal the writer's attitude’ (Hyland 1998, p.437). It encompasses ideologies from both the enunciative and *praxématique* schools, and appraisal theory, allowing for the effective identification of the author and the reader in

texts. It contains interactive resources and dimensions, and this breadth accounts for social facets of writing and reader/writer interaction at the interpersonal and textual level. Moreover, there is ample evidence to suggest its presence, albeit to varying degrees, within the academic research article in English, French, and Spanish across a range of disciplines, including economics. While there is research that does address the role of metadiscourse in the economics research article specifically, there is by no means an extensive collection of such studies. Therefore, this thesis can contribute further perspectives on interpersonal metadiscourse in economics research articles to the literature on academic writing. Hyland, who has contributed substantially to research on metadiscourse, has proposed an extensive model for studying evaluative metadiscourse in academic writing which pertains to writer stance and reader engagement. This forms the basis of the discussion in the following section.

2.9 Evaluation: writer stance and reader engagement in academic writing and the academic research article

Within the domain of interpersonal metadiscourse, the concept of evaluation and its subcategories of writer stance and reader engagement represent a growing area of interest in the field. This section elaborates on evaluative metadiscourse by outlining the current state of research on evaluation and, within that, writer stance and reader engagement. In so doing, the aim of this section is to demonstrate the importance placed on writer stance and the relative dearth of literature on reader engagement. First, the concept of evaluation is considered.

Somewhat polysemous in the literature, evaluation refers to the way writers express opinions and values, organise discourses, and create a relationship with the reader (Hunston and Thompson 2000, p.5; Mauranen and Bondi 2003, p.269). Effectively, evaluation transforms writing from what might be perceived as objective and impersonal to a rather persuasive and interpersonal enterprise (Hyland 2005a, p.185). This view sees language used ‘to acknowledge, construct and negotiate social relations’ (Hyland 2004b, p.5) and to guide the readers’ understanding of texts (Suarez Tejerina 2005, p.118). This is a view that is largely accepted by researchers in applied linguistics and academic writing who specialise in evaluation (Hunston and Thompson 2000; Shaw 2003; Martin and White 2005; Rentel 2006; Read and Carroll 2012; Jiang and Hyland 2015; Xie 2016).

Although expressing opinions and engaging readers is a notable staple of academic writing (Hunston and Thompson 2000, p.5), evaluation is, notably, quite an ‘an elusive concept’ (Mauranen and Bondi 2003, p.269) with many underlying complexities (Hood 2010, p.7). Although evaluation is complex and elusive, without exploiting the recognisable devices of evaluation to be simultaneously convincing and conformist in academic writing, it is unlikely that any contribution by a writer will be wholly accepted by their community (Hyland 2004c, p.13). For this reason it is unsurprising it has received so much attention in the literature.

Overall, in the context of academic writing, evaluation constitutes a social act in an academic ‘disciplinary or institutional context’ (Hyland 2005a, p.175) that is ‘not made from all the alternatives the language makes available, but from a more restricted sub-set of options which reveal how [writers] understand their communities through the assumptions these encode’ (Hyland 2005a, p.175). Therefore, it is valuable for academic writers to learn the appropriate use of evaluation in academic writing, and this has been a focus in research on academic writing pedagogies in English (Xie 2016), French (Rentel 2006; Tutin 2010a; 2013), and Spanish (Aguirre 2010; Lafuente-Millán 2014). Overall, the realisation of evaluative values in a text necessitates the use of specialised, frequent, and recurring rhetorical devices. These devices are fixed in nature and serve distinct evaluative purposes (Farzannia and Farnia 2016, p.3). Hyland’s model of evaluation in academic writing is informed by ‘the most important studies in this area’ (Biber 2006b, p.103), and is seen by many as the most comprehensive model of evaluation (Orta 2010, p.81; McGrath and Kuteeva 2012, p.163; Rahimivand and Kuhi 2014, p.1495). Evaluation envelops both writer stance and reader engagement and has been popular among researchers studying French and Spanish *e.g.* Rentel (2006), Salager-Meyer (2011a; 2011b), Tutin (2013), Lafuente-Millán (2014) and Xie (2016).

Writer stance refers to ‘writer orientated features of interaction’ (Hyland 2005a, p.178) and allows the writer to demonstrate attitude and realise themselves within a text. In so doing, writers can express their judgements, ‘opinions, and commitments’ (Hyland 2005a, p.176). Stance is writer-oriented (Hyland 2005a, p.178), and Hyland’s view is recognised as encompassing the individual streams of stance apparent in earlier literature (Gray and Biber 2012, p.18) which are threefold:

- first, evidentiality, the commitment the writer shows to their claims;
- second, affect, which is concerned with attitude;
- and third, presence, which refers to the extent to which the writer has manifested themselves in the text (Hyland 2008, p.7).

Following Hyland, these three agents of stance are achieved through the use of: hedges, boosters, attitude markers, and self-mentions (2005a, p.177). In brief, Hyland defines hedges as devices used to withhold commitment from a proposition, while boosters amplify a view of certainty and commitment to a proposition. Attitude markers show affective evaluations of content and self-mentions are signalled through the use of first person and possessive pronouns as well as self-reference in in-text citations.

A review of the presence of stance markers in English academic writing shows that hedges are the most commonly used, followed by attitude markers, boosters, and self-mentions (Hyland 2004c, pp.17-18). This view is supported elsewhere, where hedging appears to be very frequent in English, as English is typically not very critical (Salager-Meyer *et al.* 2003, p.223). Further studies of English find that self-mentions occur more frequently in economics than in applied linguistics research articles (Khedri *et al.* 2015, p.312), and stance nouns such as “advantage” or “difficulty” differ greatly in usage across disciplines in academic writing in research articles (Jiang and Hyland 2015, p.541). In French, hedges do occur in academic writing, but less so in the soft sciences (Vold 2006, p.69), and boosters appear to be quite frequent in French, which supports the claim that French academic writing takes a critical perspective (Salager-Meyer *et al.* 2003, p.223; Salager-Meyer 2011a, p.36). In contrast, hedges appear to be quite infrequent in Spanish (Salager-Meyer 2011a, p.36), which supports the view that Spanish behaves even more critically than French (Salager-Meyer *et al.* 2003, p.223). However, beyond such criticality, Spanish generally exhibits an overall lack of evaluative connectors (Mur-Dueñas 2007a, p.43). Spanish also appear to use many self-mentions and attitude markers in engineering papers (Aguirre 2010, p.50) while Mur-Dueñas (2007b, pp.148-149) also finds that, while English makes frequent use of self-mentions in business management research articles, this is not the case for Spanish. This indicates disciplinary differences as well as cross-linguistic differences.

Overall, there exist many studies that focus on writer stance in academic writing. However, interestingly, reader engagement has ‘been relatively neglected in the literature’ (Hyland 2005a, p.182). Addressing this gap, reader engagement is a central focus of this thesis. Nonetheless, the research on stance discussed in this section is still noteworthy since, owing to the lack of research on engagement, stance research can offer a critical mass of studies that deal with the interpersonal nature of writing on which to stage this thesis. This stance research is important in informing the discussion of the findings in Chapter 7, for example. Next, given its centrality to this study, the concept of reader engagement is discussed in greater detail.

2.9.1 Reader engagement in academic writing

Hyland has become widely accepted as the authority on reader engagement in academic writing (Biber 2006b, p.6), and he identifies that reader engagement serves two primary functions. These are to ‘include readers and meet expectations [and to] rhetorically position the audience’ (Hyland 2005a, p.182). Engagement is seen as a rhetorical tool that is crucial for writers to exploit in the effort to anticipate reader reactions (Lafuente-Millán 2014, p.202). In order to achieve this, certain markers are used to address the readers and, moreover, use the readers as a discursive tool for persuasive ends (Tse and Hyland 2006, p.775). That is to say that engaging the reader implicitly or explicitly achieves a number of useful rhetorical practices such as showing familiarity with the community, being polite, focusing the discourse, and guiding and encouraging the reader to an interpretation the writer proposes (Hyland 2005b, p.54). The particular manifestation of these markers is often specific to the genre and disciplinary discourse community (Hyland 2005b, p.54). This indicates that reader engagement practices are reflexive to a discourse community’s expectations and, therefore, can be challenging for inexperienced writers (Lafuente-Millán 2014, p.217). This has rendered reader engagement research valuable for informing academic language pedagogy (Lafuente-Millán 2014, p.220).

As a framework of analysis for reader engagement, Hyland identifies the use of the following engagement devices: directives, personal asides, shared knowledge, reader pronouns, and questions (2005a, p.117). He sees reader pronouns (you, your, we, our etc.) and personal asides as a means to meeting readers’ expectations and questions, and directives and shared knowledge as a means to position the audience. Reader pronouns

refer to the presence of pronouns that bring the reader explicitly into the text, and personal asides refer to language within the text that is used to deliver extra and tangential information to the reader. Appeals to shared knowledge exploit the joint membership of the writer and reader to their discourse community, and directives give an imperative or instruction to the reader.

Finally, questions ‘are the strategy of dialogic involvement par excellence, inviting engagement and bringing the interlocutor into an arena where they can be led to the writer’s viewpoint’ (Hyland 2002a, p.185). Building on several of his earlier papers *e.g.* Hyland (1998; 2001; 2002a; 2004a; 2004b; 2004c; 2004d), Hyland’s comprehensive framework, which was published in 2005 (2005a), outlines each of these engagement makers in detail. Harwood (2005) built on Hyland’s earlier work on pronouns and defined them as either inclusive or exclusive of readers. This designation is often ambiguous, which he argues is often a choice on the writers’ part. A more recent publication from Hyland and Jiang (2016) returns to this framework. In their paper, reader pronouns are relabelled as reader mentions, allowing for the capture of words that include readers but which are not pronouns, such as noun phrases like “the reader”, or in a multilingual study, verb inflections like *puedes* or *podemos* in Spanish, which do not require a pronoun to signal the person. In their study of academic writing in English across applied linguistics, sociology, biology, and electronic engineering, the most frequent engagement marker was directives in applied linguistics, biology, and electronic engineering, and shared knowledge in sociology. The second most frequent marker was shared knowledge in applied linguistics, biology, and electronic engineering, and directives in sociology. The remaining three markers varied in frequency across the disciplines, with questions being least used in electronic engineering and most used in sociology. Interestingly, in Hyland and Jiang’s (2016) diachronic study, it was found that reader engagement use changed overtime, largely decreasing in usage. In fact, the only increases documented were in the use of directives and questions in biology, and directives and reader mentions in electronic engineering. Otherwise there was a decline in the use of each reader engagement marker in each discipline. Building on Hyland’s work, reader engagement, unlike stance, has received relatively little attention in the literature by other researchers. That being said, there are a number of studies worth noting that have addressed reader engagement, especially those which look at reader engagement in a multilingual context.

As a rhetorical tool, engagement markers feature prominently in economics academic writing, as opposed to non-academic discourses (He and Abdul Rahim 2019). When writing, electrical engineers use appeals to shared knowledge to evoke the presentation of claims (Koutsantoni 2004). Similarly, academic writing in mathematics also makes great use of shared knowledge as well as reader mentions (McGrath and Kuteeva 2012). On reader mentions, fields like applied linguistics appear to use them more than psychology and research articles detailing quantitative studies use more directives (Hu and Cao 2015). In English and Spanish business management research articles, there is evidence of the frequent use of engagement markers, with 15.8 engagement markers per 10,000 words in English and 26.27 engagement makers per 10,000 words in Spanish (Lafuente-Millán's 2014, p.209). This indicates that Spanish may not be quite as reader-responsible as has been suggested in Section 2.2.2. In her study, Lafuente-Millán (2014, pp.205-206) applies Hyland's framework to the study of inclusive reader pronouns, directives and questions. Of the three, question use was least frequent. Other researchers such as Fløttum *et al.* (2006, p.160) have analysed directives contrastively in "let us" imperatives in English and French. Their study shows that French uses more of these directives than English or Norwegian, and economics more than linguistics or medicine. Khoutyz (2015 p.147) finds reader pronouns to be rather infrequent in Russian linguistics academic research articles, while frequent in English. Moreover, he finds that directives are more common in Russian than English, and personal asides and shared knowledge are a common feature of both languages. In his study, questions are the least frequent engagement marker and are more common in English than Russian. Considering both Fløttum *et al.* (2006) and Khoutyz (2015) in parallel, it is interesting to note that directives are less common in English than both Russian and French. However, it cannot be asserted how Russian and French compare to one another based on this, given the different data sets used. For Spanish, reader pronouns as well as directives appear to be a common feature of academic writing too (Lee and Casal 2014 pp.49-50). This is interesting when considered in terms of discussion in Section 2.2.2 which positions Spanish as not being particularly "reader-friendly" when compared to English. So too are the findings of Curry and Chambers' (2017 p.335) study of questions in linguistics research articles in English and French, where questions are found to be common in both French and English, but more common in French. This again is somewhat surprising, given the discussion of

French as reader-responsible in Section 2.2.2. This contrasts to other studies, where questions appear to be quite infrequent in Spanish academic writing (Lafuente-Millán 2014, pp.205-206). Salas (2015 pp.28-29) demonstrates that engagement in the genre of the research article in Spanish varies greatly across disciplines with regards to the manner and degree of engagement with the reader while, overall, research has shown that engagement markers are less frequent in Spanish than English (Pérez-Llantada 2010, p.47). Similarly, Lorés-Sanz *et al.*'s (2010) edited volume on interpersonality in written academic genres has shown that reader engagement varies across discipline and language, with Spanish typically making less use of engagement markers. However, their work does not deal with questions as an engagement marker.

What this brief review of research on reader engagement shows is that, while Hyland's approach was designed for and implemented in the analysis of English academic writing, his work has been adapted and appropriated beyond that context, with studies analysing academic writing in French and Spanish. His approach and theorisation of the field is well established in a life's work on metadiscourse. However, that is not to say it should be applied uncritically. While the research on reader engagement discussed above shows the applicability of Hyland's framework of reader engagement to languages other than English, and goes some way to account for its continued relevance to research in applied linguistics, there is also evidence of the need to evaluate the efficacy of the framework for the study of French and Spanish. This need is well presented by Lee (2011), who discusses comparing Japanese and English using Hyland's framework:

There are linguistic devices other than those listed in Hyland (2005), which need to be examined for expressions of stance and engagement. Cleft construction and sentential nominal are such examples.

In addition, Salas (2015, p.25) identifies issues with the application of Hyland's work on metadiscourse to languages other than English in his analysis of metadiscourse in Spanish research articles in economics, medicine, and linguistics:

Most of the studies of metadiscourse in Spanish have adopted Hyland's (2005[a]) taxonomy in its entirety and without modification. This can be brought up as a disadvantage due to the fact that it was proposed for English and it should not be assumed that the same metadiscursive features exist in Spanish.

In the analysis of question functions in English, French, and Spanish, this thesis makes use of a modified version of Hyland's framework of question functions, which forms part

of his framework of reader engagement. The decision to use a modified version of Hyland's framework was not made lightly. Hyland's framework of reader engagement and questions very much sets the architecture for the identification of question functions. However, since the development of his framework, others have contributed to the field allowing for a more enriched framework to be developed. The decision to use the modified version of Hyland's framework emerged from its successful application in a pilot study wherein, owing to its breadth, it captured all the question functions identified in the data. These were categorised based not solely on Hyland's work but on a synthesis of literature on question functions, discussed in Section 2.10.1. Bearing this enrichment in mind, Hyland's framework is suitable for the analysis of English, French, and Spanish, in so far as it accounts for all question functions identified in KIAP-EEFS.

It should be noted too that the convergent nature of the contrastive study presented in this thesis renders the use of Hyland's framework less problematic. Convergent and divergent contrastive studies are discussed in detail in Chapter 3, Section 3.3.3. In brief, taking a convergent approach means that the focus from the outset is on questions, which are identified independently in each language and then compared across languages. In contrast, a divergent study of questions would identify questions in English only, and then compare these questions to their equivalences in French and Spanish, which may include questions but also other structures. Were this study a divergent study, Hyland's framework would be less suited, as a divergent study could necessitate the equating of questions in English to engagement makers in French and Spanish that do not exist within Hyland's framework.

Overall, much less research has centred on reader engagement as evaluation, with many researchers focusing instead on the writers' stance as opposed to the role of the reader (Hyland 2005a, p.182). In terms of reader engagement, of particular note are questions, which are under-represented in research, especially in French and Spanish, despite being recognised as one of the most direct means of overtly interacting with the reader (Thompson 2001a, p.59). To address this gap, this thesis contributes to research on questions.

2.10 Questions as reader engagement in academic writing and the academic research article

Questions as reader engagement form the basis of this contrastive research. This section presents an overview of research on questions used to engage readers, first in terms of question function in Section 2.10.1, and second in terms of question type and form in Section 2.10.2.

2.10.1 Question functions

As mentioned in Section 2.9.1, this thesis adopts a modified version of Hyland's framework of question functions as the broad architecture for question function in this study. In terms of reader engagement in academic writing, questions have had rather little attention in English, French, and Spanish, although Hyland has made a strong contribution to this field in relation to English. This section first discusses Hyland's framework of seven question functions to set the backdrop on which this study bases its understanding of question functions as reader engagement devices. Following that, further literature on the study of questions are discussed, which serve to complement Hyland's framework; enrich the functional perspective of questions; and offer a more complete view of questions and their function as engagement markers in the academic research article. Overall, the aim of this section is to summarise what research has been conducted on question functions in questions used as reader engagement, and outline the conceptual framework used to categorise questions according to function which, as presented in Chapter 1, Section 1.4, is a core research question addressed by this thesis.

Questions have long been of interest to research in academic discourse, especially from a functional perspective. For example, Webber (1994) takes an affective view of questions, seeing their function as creating anticipation, interest, and thought in academic writing, clearly serving an evaluative purpose. Her analysis of medical papers in English sees questions used to create distance and hedge, close discussions and point forward, and provoke and criticise. She summarises the evaluative and engaging role of questions as follows (1994, p.266):

The most important objective of using interrogatives in the corpus studied is that of reader involvement. Questions create anticipation, arouse interest, challenge the reader into thinking about the topic of the text, and have a direct appeal in bringing the second person into a kind of dialogue with the writer, which other rhetorical devices do not

have to the same extent. At times, they may be discreet. At other times, they may be very aggressive when they are used to attack opponents as part of academic debate.

Elsewhere, questions have been referred to as ‘inappropriate’ (Swales and Feak 1994, p.74) in academic writing. Nonetheless, within the realm of evaluation and academic writing, questions have been seen as one of the primary means of overtly interacting with the reader (Thompson 2001a, p.59), where they enter, more than any other engagement marker, into direct involvement with the reader (Thompson 2001a, p.59) in an effort to make readers follow the writer’s line of questioning as ‘reader-in-text’ (Thompson 2001a, pp.59-60). Hyland (2001, p.170) posits that questions are a rhetorical positioning and engagement of the reader used to make the reader feel like a judge whilst offering them no opportunity for response (Hyland 2001, p.570). Questions, he argues, can also be used to close articles dramatically, in order to maintain the readers’ interest (2001, p.570). Moreover, Hyland has found that the use of questions in academic writing varies across disciplines, where questions are more common in the soft sciences than the hard sciences (2002a, p.537). Hyland’s framework of question functions is the most comprehensive in the literature, and systematically categorises questions according to seven broad functions, namely, getting attention; framing the discourse; organising the text; creating a niche; expressing an attitude and counter-claiming; setting up claims; and real questions (Hyland 2002a, p.535). These functions are now discussed.

Beginning with getting attention, Hyland identifies that the ‘title is generally the reader’s first encounter with a text and may be the point where a decision is made whether to read on and give the accompanying paper further attention, or to ignore it’ (Hyland 2002a, p.539). He continues that questions in titles ‘can help grab the reader at the outset with an arresting directness’ (Hyland 2002a, p.539). This is corroborated by researchers such as Buter and van Raan (2011, p.616) and Cook and Plourde (2016, p.1120), who argue that publications with questions in their titles are more cited than those without. Essentially, for Hyland, questions that get readers’ attention in academic publication titles serve to hook them, and are ‘never a naïve display of ignorance, nor do they simply signal the topic of the paper’ (Hyland 2002a, p.540).

Next, in defining questions that frame the discourse Hyland (2002a, pp.540-541) identifies that:

in a traditional sense, academic writing is governed by questioning. It is written with a question in mind, a problem to solve, and seeks to explore that question using the theories and procedures of the discipline. It is perhaps not surprising that these questions often surface in introductions to provide an initial framework for the discourse [with] writers posing the questions to be addressed [...] through the text to hold the reader's interest and structure their responses to them.

For example, framing the discourse largely refers to questions that are posed in order to fill a gap in the literature. They are the questions that the research article systematically intends to answer. In this way, questions that frame this discourse often serve to deal with the research question and sub-questions that guide the research article. Moreover, they are questions that cannot be answered right away. This is quite distinct from questions that organise the text, which fulfil a more structural role.

According to Hyland (2002a, p.544), in organising the text, questions occur in subheadings, which indicate the content of the following section to the reader. As a signalling tool, they direct the reader to an argument in the upcoming section that is focused on answering the question posed in the subheading; however, these questions are not necessarily research questions or sub-questions, but are questions that allow the writer to move the argument in a certain direction, serving a signposting and signalling function. Within the text, questions that serve to identify the end of one discussion and the start of another are considered text organising. They guide the reader's thinking and questioning in the same direction as that of the author. This allows the readers to navigate the text more easily and encourages less rejection or objection on their part. Hyland's (2002a, p.544) examples illustrate both explicit use of the metalanguage of text organisation, and also 'the use of in-text signals to navigate the reader through an argument'. The following examples, taken from Hyland (2002a, p.544), illustrate these uses respectively:

What are we to conclude from these data?
How should one express gratitude for a meal in another culture? Is it possible to refuse an invitation politely? How should one greet people in different speech communities?

The fourth function of questions that Hyland presents is creating a niche. This term is borrowed from Swales (1990). Hyland (2002a, p.545) defines introductory questions as serving a:

rhetoical strategy for 'establishing a niche' for one's research. More specifically, writers are able to use this strategy to raise questions about current theories [...], or to indicate a gap in the literature [...]. In addition to establishing a niche, research writers also used questions to claim centrality for their work, a strategy which appeals to readers 'to accept that the research

about to be reported is part of a lively, significant or well-established research area' (Swales 1990: 144).

Therefore, questions that create a niche are unlike discourse framers, as they do not refer to the area of research, the architectural structure of the research article, or the questions to be answered over time. Rather, they allow the author to, firstly, identify a gap by raising questions that have not yet been answered and, secondly, to claim centrality. They claim centrality by asking questions of the literature that only they can move to answer, which allows them to generate novelty.

The function expressing an attitude and counter-claiming is rhetorically linked to the function creating a niche; however, its primary difference is its use 'to express an evaluation of an idea, either positively or negatively' (Hyland 2002a, p.546). More specifically, Hyland (2002a, p.546) argues that, when questions are used in this way:

This is frequently a critical posture, or at least one which carries a strong intimation of distancing the writer from the proposition and is usually followed by a reason for the questioning. [In such questions] the reader is recruited into a virtual debate to support the author's evaluation through the question. As readers we are treated both as knowledgeable colleagues and active participants who share the writer's concerns. This strategy, moreover, is not only used to critically examine others' work, but to present tentative counterclaims by framing what appear to be the reader's reservations regarding the validity of a foregoing argument.

Thus, such questions are identifiable through both their critique of others' work and the proposal, on the authors' part, of another means by which something could be understood. In such a question, what becomes important, as in the case of each function identified thus far, is the importance of cotext and context. That is to say that, in most situations, it is not the question, but the subsequent proposition and its enveloping sentences, context, and purpose, that help to define the question's function.

Hyland's sixth function of questions is their use in setting up claims. Unlike expressing an attitude and counter-claiming, setting up claims does not require that the author questions or critiques another's work. Rather, questions that set up claims are (Hyland 2002a, p.548):

questions that [...] play an important role in presenting the writer's stance, challenging the reader to think about an issue, but anticipating the assertion of an opinion.

Such questions are underpinned by the basic understanding that questions, in the interactive research articles, differ from conversational questions in that, it is unlikely

that, in a finalised research article, questions ‘originate in ignorance [and are] motivated by the writer’s desire to gain information or to think aloud about some unsolved mystery’ (Hyland 2002a, p.548). Rather, Hyland argues that, for the express purpose of setting up claims (Hyland 2002a, p.548):

Their importance lies in helping the writer to create what appears to be a jointly constructed textual environment for exploration, providing writers with a rhetorical space to contribute to the completion of the text and the achievement of its objectives.

Finally, questions that ask real questions and point forward constitute Hyland’s seventh question function. Such questions ‘genuinely [seek] information rather than [anticipate] a specific response’ (2002a, p.552). Imaginably, real questions typically occur in conclusion sections of soft science research articles (Hyland 2002a p.552), where they raise ‘unresolved issues and emergent matters to be addressed in future studies’ (Hyland 2002a p.552). To summarise the role of such questions, Hyland (2002a p.552) refers to Smith (1987, pp.19-20):

It is common for a piece of research to answer the question it has set out to clarify while at the same time it raises other questions to be accounted for in the course of further investigation.

Thus, although appearing more akin to questions from conversational contexts, such questions also serve a rhetorical purpose. That is, they allow writers to raise questions to which they want an answer, and simultaneously position themselves as researchers who can identify potential issues for future research – a quality arguably within the remit of a good academic and scientific researcher (Smith 1987, pp.19-20; Hyland 2002a, p.552).

While Hyland’s framework of the seven functions of questions is both broad and comprehensive, since its conception in 2002, further research on questions has added to understandings of question functions. Questions in academic writing are viewed as opportunities to ‘hook’ (Ruegg and Sugiyama 2013, p.13) readers, and are expressly used for the purpose of interaction (Hyland 2005b, p.537). Such a hook further explains the choice to use questions that get attention, for example. However, the function of hooking readers can also help deepen our understanding of Hyland’s real questions and those that create a niche. Others who have considered questions have seen them as, for example, the introduction to ‘click-bait’ (Hamby 2015, p.107) in academic writing, where questions are used to make interesting titles (Cook and Plourde 2016, p.1120). Furthermore, Ball (2009, p.677) studies questions in titles and finds that they are more frequent in titles in

the natural sciences, especially medicine. He argues that this has come about because of the need for visibility and engagement in academia for career development. Questions, he argues, help focus the reader on the topic of the paper directly from the title, which lends a further depth to questions in titles that, in Hyland's view, mainly serve to get the reader's attention. For Nagano (2010, pp.101-102), questions occur more often in titles in soft sciences and are more frequent in economics than history or sociology. These questions, she argues, are typically full sentences and appear in the first unit of a two-unit title. Interestingly, Buter and van Raan (2011, p.616) identify that the use of questions in titles of publications sees a higher degree of citations for those publications, which Cook and Plourde (2016, p.1120) imply can be a motivation for writers to employ the use of questions. Others, like Soler (2011, p.133), find questions useful for re-opening discussions or emphasising, reflecting earlier studies pre-dating Hyland, such as Webber (1994).

Beyond titles, there is much less research on questions in academic writing. Researchers such as Pic and Furmaniak (2014, p.370) see questions as a means to question certainty in scientific and popular science research articles. Furthermore, they identify that authors can use question marks to protect themselves, by expressing 'the dubious nature of results'. This view builds on Hyland's function of setting up claims, for example, as such questions include the reader in the construction of knowledge, where the writer feels more tentative about a claim. Overall, for Soler (2009, p.57), questions in academic writing indicate the anticipated content of the paper and fulfil a rhetorical function, a rhetorical function summarised well by Soler (2009, p.57):

Cuando el objeto de estudio se presenta con estructura de pregunta, el investigador/autor parece perseguir alguno de los siguientes objetivos: i) autopreguntarse sobre el objeto de estudio y encontrar a través del trabajo que reporta en su artículo, una respuesta para su interrogante; ii) evitar una aseveración directa porque hay aun interrogantes sin respuesta; o iii) invitar al lector a que lea el artículo, despertando así su interés por un determinado problema. En todos estos casos hay interacción entre el autor y el lector. Aun en la instancia i), el propio autor estaría cumpliendo en una sola persona el rol de ambos participantes en todo acto de pregunta.

Others, like Ryvitytė (2003, p.95), find that in move 2 in economics research articles, there are more examples of question raising for identifying a niche, supporting Hyland's function of creating a niche. Moreover, she finds this to be especially the case in economics when compared to linguistics or medicine. Of note too is a previous study applying Hyland's functions to questions in linguistics in English and French by Curry

and Chambers (2017). In that study, English and French both make most use of questions mainly to organise the text and set up claims.

Reflecting on the functions of questions discussed in the first part of this section, Hyland's seven functions broadly capture a range of roles that questions perform. Building on his work, further studies have enriched the understanding of the function of questions as reader engagement. In Section 2.9, reader engagement and the importance of reflecting on the relevance of Hyland's work to languages other than English were discussed. Once again, it is worth restating that the decision to use a modified version of Hyland's framework emerged from its successful application in a pilot study, wherein modified version of Hyland's framework captured all the question functions identified in the data. The first modification refers to Hyland's first function, getting attention. This function has been broadened to see questions in titles as not only functioning to get readers' attention, but to focus the topic of the research article from the outset. Therefore, in this study, it is referred to as "getting attention and focussing the reader", given the evidence discussed above that questions in titles are more than just interest raisers. The second modification pertains to Hyland's sixth function, setting up claims. Here, recognising the value of questions used to protect writers from their dubious or tentative claims, setting up claims as a function is broadened to "setting up claims and protecting the writer". This refers to the writers' use of the reader when asking questions in order to set up claims. Writers allow readers to be co-constructors of the problem, and validators of the answer, through questioning. This shares the burden of the claim and protects writers in the context of less certain claims. Moving on from this discussion of the function of questions as reader engagement devices in academic writing, the following section discusses the types and forms of questions studied herein.

2.10.2 Question types and question forms

When searching for questions in academic writing, their presence is not always so clear. Questions are signalled in writing by 'illocutionary force indicating devices' (IFID) (Flöck and Geluykens 2015, p.7) which are 'any element of natural language which can be literally used to indicate that an utterance of a sentence containing that element has a certain illocutionary force or range of illocutionary forces' (Searle *et al.* 1985, p.2). While most research on IFIDs has centred on spoken language, its relevance to academic

research rests on the view of academic writing as interpersonal and dialogic (Blagojević and Misić-Ilić, 2012, p.17), and as pertaining to a discourse, where illocutionary acts serve as a social form of linguistic behaviour, that exploit propositional content to denote an intended purpose (Vanderveken 2001, p.243). As Section 2.8 and Section 2.9 have shown, the interactive and interpersonal nature of academic writing is very well established. Furthermore, following the discussion in Section 2.10.1, the interactive and interpersonal nature of questions in academic writing serving to perform myriad functions demonstrates the range of illocutionary acts questions can perform in academic writing. Here, it is not possible to go into detail on the theorisation of IFIDs as has been done elsewhere – see Searle *et al.* (1985) and Krifka (2014). Rather, the concept of IFID is important for finding both direct and indirect questions in a corpus (Karttunen 1977, p.4).

Direct questions are identified by the IFID of a question mark rendering them easily corpus searchable. Karttunen, (1977, pp.4-5) proposes that indirect questions are signalled with IFIDs such as *wh-* clauses which are clauses that contain question embedding verbs followed by words like “whether”. Table 2-1 shows the question embedding verbs that function as IFIDs which are used to raise indirect questions as identified by Karttunen (1977, pp.384-385).

Table 2-1 Verbs used to embed questions in Karttunen (1977, pp.4-5)

Verbs of retaining knowledge	know, be aware, recall, remember, forget
Verbs of acquiring knowledge	learn, notice, find out, discover
Verbs of communication	tell, show, indicate, inform, disclose
Decision verbs	decide, determine, specify, agree on, control
Verbs of conjecture	guess, predict, bet on, estimate
Opinion verbs	be certain about, have an idea about, be convinced about
Inquisitive verbs	ask, wonder, investigate, be interested in
Verbs of relevance	matter, be relevant, be important, care, be significant

Verbs of dependency	depend on, be related to, have an influence on, be a function of, make a difference to
---------------------	--

Romero (2005, pp.687-688) builds on this work, proposing ‘concealed questions’ as a type of indirect question, where sentences that do not hold typical question or embedded question syntax correspond to embedded interrogative clauses. For example, he proposes that the following are both indirect questions:

John knows Bill’s telephone number.

John knows what Bill’s telephone number is

He sees that the first example is equal to the second in terms of its questioning capacity. However, this is only the case when the intention is to reveal some information *i.e.* John knows Bill’s telephone number and John intends to tell us what it is. Therefore, in this view, indirect questions do not need to have a *wh-* complement, but are defined by their relationship to the propositional content. As such, indirect questions become more difficult to identify as the IFIDs are not necessarily linked to a *wh-* complement. The following examples show a *wh-* complement indirect question in English and one without in Spanish from KIAP-EEFS:

It is not clear whether inflation rises or falls with the amount of time left in office.
(engecon27)

Es difícil saber si ha sido un error estratégico o una cesión inevitable ante el poderoso empuje de la nueva Alemania (specon15)

These questions are raised with the intention of revealing some information, and these same types of questions have been documented in French (Hirschbühler and Valois 1992) and Spanish (Plann 1982). Seeing as indirect questions with or without a *wh-* complement can share meaning with direct questions, it is reasonable to accept the view that Suñer (1993, p.46) proposes, whereby ‘interrogatives and *wh-* complements are to be analysed in terms of a uniform semantic object’. That is to say that each indirect question relates to a direct question form, and direct and indirect questions can be categorised in the same ways. Reflecting upon this first binary of question types as direct or indirect, there emerges an evident challenge in the identification of IFIDs that generate indirect questions. For those indirect questions studied here, Chapter 4, Section 4.2.3 gives a step-by-step description of the process for identifying direct and indirect questions in KIAP-EEFS. These questions take different forms depending on their syntactic make-up. Given

the range of modes, genres, and texts in which questions are used, for brevity, the question forms found in this study are identified and discussed only.

Following Karttunen (1977, pp.5-6), wh- questions, yes/no questions, and alternative questions are important here. Wh- questions are those that require content and information as a response. Following San Roque (2016, p.85), in this study, these are called content questions so as not to privilege English typology of wh-. The following are examples of content questions in KIAP-EEFS:

Why are the magnitudes of the gains based upon these two approaches so different?
(engecon06)

Quel arbitrage efficacité-équité, dans le financement et la gestion de la filière de soins, chacune de ces deux conceptions du réseau implique-t-elle? (frecon03)

¿Cómo podemos inmunizar contra este tipo de riesgo el valor presente de nuestros flujos esperados? (specon20)

Karttunen (1977, p.5) also identifies yes/no questions as questions which necessitate a yes or a no as a response. Nagano (2010) also studies yes/no questions, and Farkas and Bruce (2010) label these polar questions, which again, as a label, better fit this cross-linguistic study. Dietz (2001) finds that titles in questions most often make use of complete polar questions. The following are examples of polar questions in KIAP-EEFS:

Une banque paie-t-elle plus cher une offre accrue de certificats de dépôts ? (frecon48)

¿Están mal remunerados los profesores? (specon44)

Karttunen (1977, p.5) also refers to alternative questions, which are questions in which there is some element of choice in, as in:

Does firm size or higher profitability explain the impact of foreign investment?
(engecon41)

La « normale » désigne-t-elle un niveau moyen de stocks, constant au cours du temps, ou bien un niveau de stocks variable, dépendant de la conjoncture du moment?
(frecon40)

¿Tiene este en los años 2000 más o menos la significación de las transacciones de los años 1900 o por el contrario tal significado se ha visto modificado sustancialmente?
(specon13)

Borrowing from Carter and McCarthy (2006, p.181), Curry and Chambers (2017) identify declarative questions as questions that form a syntactically complete affirmative statement, where a question mark is tagged on at the end, as in:

[...] nous nous demandons quel partage des risques en consommation est atteint par les ménages ruraux au Pakistan et quels marchés « d'assurance » leur permettent de s'assurer complètement ou partiellement? (frecon29)

Finally, Bloom *et al.* (1982, p.1088) identify elliptical questions. Elliptical questions are those that are not syntactically complete. They have elided certain words that would traditionally be integral to question formation *e.g.*:

Macroeconomic fluctuations: Demand or supply, permanent or temporary?
(engecon22)

Trop ou trop peu de produits en concurrence? (frecon46)

¿La causa? (specon22)

Elliptical questions, however are a somewhat controversial designation. Ellipsis is a syntactic feature and it could be argued that any type of questions mentioned already could be ellipted. Is ¿La causa? for example, simply an ellipted declarative question? Ellipsis appears to be a common feature of spoken discourse, and its presence in question formation in academic writing is worth capturing. As such, it was decided to capture elliptical questions as a distinct form of questions in order to set them in relief.

Overall, these same question types are reflected in French and Spanish grammars (Bradley and Mackenzie 2004; Pountain and Kattán-Ibarra 2004; Hawkins and Towell 2015); however, there is scant research on question forms in academic writing in French and Spanish.

In Nagano (2010, pp.109-110), content, polar, and alternative questions occur, with polar questions most common in economics. Biber *et al.* (1999, p.214) find questions to be infrequent in academic writing, and where they do occur, 55% are content questions and 40% are polar questions. The final 5% is composed of alternative questions, declarative questions, and tag questions. In Blagojević and Misic-Ilic (2012, pp.22-23), the English academic articles are seen to contain content, polar, and negative questions, as well as embedded questions. They identify embedded questions, also known as indirect questions, through the identification of IFIDs in the form of verbs, nouns, and other parts of speech used to raise questions. Finally, Curry and Chambers' (2017, p.336) study of linguistics research articles shows that French contains mainly content, polar, and declarative questions while English contains mainly content and polar questions. Research investigating question forms is important, as researchers such as Beun (2000,

p.311) have indicated that different question forms are more likely to occur within different registers and modes such as spoken and written registers. For example, according to Beun (2000, p.311), declarative questions are associated with spoken discourse, at least in the case of English. This is interesting as there is evidence of declarative questions as a feature of French linguistics articles in Curry and Chambers (2017, p.337). This shows a clear difference between English and French questions in academic writing. Overall while there is little research on questions and their form in academic writing, the situation is that much graver in the context of languages other than English, where there is scant research on questions in academic writing in French and Spanish, with the few studies addressing questions in terms of function and form peripherally across languages addressed here throughout Sections 2.10.1 and 2.10.2.

Before closing this chapter, it is worth reflecting on a key study that led to the conceptualisation of this thesis: Curry and Chambers (2017). Their study analysed questions in academic writing in the linguistics research article in English and French using the KIAP corpus. In that study, it was found that, in terms of functions, English and French use questions similarly. However, the extent to which specific question functions are used differs considerably. This is especially true for questions functioning as discourse framers and text organisers, which represent the most diverse and most used question types respectively. When considering the form of these two functions there is a tendency towards more yes/no questions in French than in English, and a much larger use of what they group as “other questions” in French than in English. Moreover, where “other questions” occur, there seems to be a greater use of ‘alternative questions’ in English and a lack of ‘declarative questions’, which are most common in French. Overall, however, there are also strong similarities in the high occurrence of content questions in both English and French for both functions. In terms of location, the English data largely corroborate Hyland’s research (2002a, p.541) which sees questions that frame the discourse occur largely in the introduction sections. However, in French, questions that frame the discourse are quite often used in the thematic middle section. This is much less common in English. Moreover, questions used to organise the text are distributed relatively equivalently across both English and French.

Overall, Curry and Chambers' (2017) study, while very much a pilot, inspired this more in-depth analysis of reader engagement as well as the concurrent analysis of Spanish. Questions as reader engagement are under-researched, especially in languages other than English. Moreover, almost no studies in reader engagement consider indirect as well as direct questions, and no other studies found in the literature have contrastively analysed questions in terms of frequency, function, question length, question form and type, question location, use of passivity in questions, use of tense and verbal modality in questions, and question sentence types in three languages as this study has. As such, the overall dearth of research on questions as reader engagement in academic writing in English, French, and Spanish is a gap this thesis addresses.

2.11 Conclusion

This chapter has focused on academic language in the multilingual context, discipline and genre, academic writing, metadiscourse, reader engagement, and questions in the (economics) academic research article in English, French, and Spanish. In Section 2.2, the global context of academic discourse was discussed. There, the role of English as the lingua franca of academia and the role of French and Spanish and other academic languages were juxtaposed, in order to problematise the comparability of academic discourse in any language with English. The impact English as the lingua franca of academia on other languages was discussed, focussing on the risks of domain loss and diglossia inherent in the expansion of English within global academic domains. Following that, the impact of culture and academic tradition was discussed, highlighting how language is an important variable when studying academic discourse across languages. Centring on concepts of reader-responsible and writer-responsible languages, as well as content-oriented and formal-oriented languages, English, French, and Spanish academic discourse was discussed and positioned within these paradigms. Overall, Section 2.2 contextualised this research in a multilingual context and signalled the importance of research on languages other than English. Moreover, Section 2.2 served to problematise the comparability of academic discourse across languages and with English.

To address this problematisation, Section 2.3 discussed the concept of discipline and the disciplinary discourse community. In so doing, Section 2.3 positioned the data in KIAP-EEFS, which are outlined in detail in Section 4.2, as including the language of three

separate disciplinary discourse communities in economics in English, French, and Spanish and one focus disciplinary discourse community in economics, spanning language, space, and time. Having established the discourse community studied herein, the focus on the text it produces was then considered, which led to a discussion of genre in Section 2.4. The discussion of genre in Section 2.4 served to ensure the comparability of the research article in English, French, and Spanish. Understanding the textual, discursive, rhetorical, and social nature of the genre, as well as its designated readership, allowed for its conceptualisation as a disciplinary genre that belongs to the focus disciplinary discourse community. Establishing the genre of the research article being studied here, Sections 2.5-2.7 discuss the textual and linguistic features of the academic research article in English, French, and Spanish respectively. These sections each discussed the structure of research articles, the length of sentences and sentence types used therein, and the role of passivity, tense, and verbal modality in research articles in each language. As these sections highlighted the textual features of academic research articles, the subsequent section moved on to discuss the contextual features in terms of metadiscourse in Section 2.8. The role of metadiscourse was discussed to highlight the importance of the interactive nature of the research article, with a perspective that sees the interaction as one between the writer and the reader. In Section 2.9, the evaluation framework used in this study was discussed, focussing largely on the importance and lack of studies on reader engagement. In that section, the gap in the literature on reader engagement, especially in a multilingual context, was identified. This gap pertained not only to reader engagement research, but specifically to the role of questions in facilitating reader engagement, which forms the basis of the discussion in Section 2.10. In Section 2.10, questions and their function as devices of reader engagement and form in written language are discussed. The discussion draws attention to the lack of research on the topic in the field, a dearth which is even greater in research French and Spanish academic writing. Therefore, Section 2.10 signals the important contribution this thesis makes to research on questions. Having established the domains of research to which this thesis contributes, there remains a further body of knowledge worth discussing. That is the role of contrastive linguistics and corpus linguistics in informing the data collection, analysis, and interpretation of data which are discussed in the following chapter.

Chapter Three

Corpus-based contrastive analysis:
background, theoretical foundation,
and application to academic writing

3 Corpus-based contrastive analysis: background, theoretical foundation, and application to academic writing

3.1 Introduction

In a world of increasing ‘interlingual and intercultural communication’, corpus-based contrastive analysis is notably experiencing a second rebirth in interest which has largely been influenced by advances in the theoretical foundations of contrastive and corpus linguistics (Granger 2003, p.18). This chapter details these advances in order to present the current state of the art of corpus-based contrastive linguistics and discuss the theoretical and methodological foundations that have informed this research. Therefore, Section 3.2 discusses contrastive linguistics. The purpose of this section is to outline the current state of contrastive linguistics in order to identify the core theoretical concepts integral to effective contrastive studies. Focussing initially on its historical context and the concerns surrounding the relevance of contrastive linguistics to contemporary research, this sections moves on to discuss its evolution and theorisation as a field of study. Following that, Section 3.3 focuses in detail on the two core concepts in contrastive linguistics: the *tertium comparationis* and equivalence. This section discusses these complex concepts in detail, outlining their role in contrastive studies and their importance for overcoming methodological weaknesses in comparing across languages. Building on this theorisation, Section 3.4 discusses theoretical and conceptual frameworks in contrastive linguistics literature in order to position this study within a contrastive linguistic framework and illustrate how the *tertium comparationis* and equivalence work in practice. Following that, Section 3.5 moves to a discussion of corpus linguistics. The role of this section is to outline the field of study that is corpus linguistics, discuss its relevant methodological and theoretical processes and applications, and make clear its value for contrastive studies. Section 3.6 moves to bridge these two fields of study in a discussion of the contributions of this research to the field of corpus-based contrastive analysis. This section focuses on corpus-based contrastive studies of academic discourse, highlighting the studies in the area that have informed this thesis and the gaps in the field to which this corpus-based contrastive analysis contributes. Finally, Section 3.7 offers a

brief conclusion, positioning this research at the forefront of corpus-based contrastive analysis research.

3.2 Contrastive linguistics: background and current state of the art

Contemporary research in contrastive linguistics is largely underrepresented in applied linguistics literature; this however, is not without reason. Contrastive linguistics came to prominence in the post-World War II era (Griffiths and Parr 2001, p.247) and this was a time of experimentation and instability in linguistic and language education research which was experiencing ‘changing winds and shifting sands’ (Marckwardt 1972, p.7). These shifting sands led to the temporary abandonment of contrastive analysis and its controversial reclassification of as a branch of linguistic typology in the 1980s by Hawkins (1986) (Kortmann and Meyer 1992; Rohdenburg 1992). Its rejection left a stain on the reputation of the field of study which is often discredited in contemporary discourses in applied linguistics (Al-Sibai 2004). Nonetheless, contrastive linguistics has remerged as an important field of study where language pairs have unwaveringly been and continue to be analysed with a view to ‘identify the similarities and differences between the conceptual structures of different languages’ (Goddard and Wierzbicka 2008, p.206). With this in mind, this section discusses the historical context of contrastive analysis and outlines why it fell from favour. Following that, contemporary perspectives on contrastive linguistics are discussed, detailing the evolution of the field and its current state of the art in order to demonstrate the importance of the field for ensuring faithful comparisons across languages.

The contrastive analysis hypothesis proposed by Fries (1945) and developed by Lado (1957) endeavoured to compare languages for pedagogical application. Fries highlighted that for language teaching ‘the most efficient materials are those based on a scientific description of the language to be learned, carefully compared with a parallel description of the native language of the learner’ (1945, p.9). His approach became popular during the 1960s following the application of Lado’s microlinguistic contrastive analysis in *Linguistics Across Cultures* (1957). Lado’s contrastive analysis hypothesis posits that features of foreign languages similar to those in a learner’s mother tongue are easier for a learner to learn and that conversely, those features that are quite distinct from those in the learner’s mother tongue are more difficult for a learner to learn (1957, p.7). Therefore,

it was thought that errors and interference could become predictable in the language learning process if more attention were paid to the learners' first language (1957, p.7). However, the hypothesis eventually fell from favour where researchers such as Klein (1986, p.26) began to recognise the limitations of the contrastive analysis hypothesis in explaining errors.

The model of the contrastive hypothesis at this time existed in three forms: strong, moderate, and weak. The strongest form focused on error prediction (Wardhaugh 1970, p.3). The moderate form was more concerned with categorising patterns of similarities and differences between languages to inform learning (Oller and Ziahosseiny 1970, p.184), and its weak form was more removed and offered a linguistic description (Wardhaugh 1970, p.8). As researchers were moving away from contrastive analysis, the loss of confidence in its strong and moderate forms led scholars to search for new explanations within the then growing fields of error analysis and interlanguage studies when informing language pedagogies (Johansson 2007, p.2). Moreover, contrastive analysis was seen to prioritise semantics and language as *langue*. This reflects the Chomskian views that language should be understood in terms of speaker competence and not performance. Spoken language or performance, in a Chomskian view, is seen as a poor representation of speakers' abilities as opposed to competence, which refers to the latent linguistic knowledge of the idealised speaker (Chomsky 2009, p.97). This view became incongruent with growing interests at the time in pragmatics and research into language in context (James 1980, p.98). Overall, contrastive analysis was seen to be too optimistic (Sridhar 1975) and theoretically weak (Selinker 1972), and this resulted in an arguably premature abandonment.

Owing to this abandonment, interest in the field of contrastive linguistics diminished between the 1960s and 1980s when it slowly remerged in a pattern of 'success-decline-success' (Granger 2003, p.13). Its rebirth as contrastive linguistics, a theoretically sound discipline in its own right, was largely indebted to the improved theorisation of the field and technological advances in concurrent fields like corpus linguistics – for more detail on the value of corpus linguistics to contrastive linguistics, see Section 3.5.4.

Nowadays, contrastive analysis does not claim to answer all questions surrounding the best means to inform language teaching and learning. Its aforementioned weak form,

which is largely concerned with linguistic description, better conceptualises the role contrastive analysis plays in contemporary research, given that in the field of language teaching and learning, studies of texts alone are inadequate for making informed pedagogical decisions. Contemporary contrastive analysis is concerned with language as *parole* not *langue* (Chesterman 1998, p.28). *Parole* refers to the consideration of language texts within their contexts – a concept that is important to this study, as discussed in Chapter 2, Section 2.4. The reason for this change is that the study of language has undergone considerable diversification since contrastive analysis' conception (Simpson 2011, pp.1-2). More specifically, the perspective of contrastive linguistics has been redrawn following global movements towards the study of language in context *i.e.* pragmatics (Senft 2014, p.133). Pragmatics became popular due to seminal studies by linguists such as Leech (1983), Levinson (1983), and Sacks (1984) and earlier north-American research that heavily influenced the linguistic movements at the time such as Garfinkel's work on ethnomethodology (1967), Searle's work on speech acts (1969) or Goffman's work on social interaction (1971). Effectively, these studies brought about a departure from Chomskian views of language where language as *parole* was seen as a 'tool [used] to interact socially and communicate' within specific contexts and therefore worthy of analysis in its own right (Senft 2014, pp.2-3).

With the advent of pragmatics, contrastive analysis also departed from its focus on *langue*, language systems, and global syntax (James 1980, p.61) and turned to focus on *parole* (James 1980, p.98). This was quite an important turning point for contrastive linguistics as, although as an approach it has existed since at least the 9th century (Huise 1624; Eynsham c. 955 - c.1010), language as *parole* has only been a feature of contrastive linguistics since James' work on approaches to contrastive analysis (1980; Chesterman 1998, p.28). James (1980) and later Krzeszowski (1990) and Chesterman (1998) see a large overlap in the distinctive approaches in both contrastive linguistics and translation studies which look at language in its wider socio-discursive context. Following the discussion of the disciplinary discourse community in Section 2.3 of Chapter 2 and the interactive nature of the research article in Sections 2.8-2.10, it has been made clear already that the consideration of the wider discursive context is critical to this study. The theorisation of the field of contrastive analysis, however, did not stop at a movement towards the study of language in context. Further consideration was given to core

underpinning concepts such as the *tertium comparationis* and equivalence. These concepts are important as they help ensure effective and faithful comparisons across languages. Throughout Chapter 2, issues of comparability were raised. Section 2.2 queried the possibility of faithfully comparing English with other languages. Section 2.3 interrogated the comparability of the discipline of economics across languages and Section 2.4 considered the comparability of the economics research article across languages. These questions are integral to ensuring a comparable *tertium comparationis* and analysing it in terms of its equivalences. The following section discusses these two concepts in detail.

3.3 Theoretical foundations in contrastive linguistics

The previous section discussed the historical context and evolution of contrastive linguistics in order to address the misconception of contrastive linguistics as theoretically lacking. Building on this view of the current state of the art, there remain some concepts of importance worthy of discussion. The emerging interest in language as *parole* led contrastive linguistics to converge, at least theoretically, with translation studies. This convergence occurred as these two different fields of study both hold at their core two complex concepts: the *tertium comparationis* and equivalence. Both the *tertium comparationis* and equivalence play an integral role in contrastive linguistic approaches and methodologies and this has been the case throughout the evolution of the field, albeit to varying degrees. The following sections consider these two core concepts.

In an effort to simplify the rather complex concepts of *tertium comparationis* and equivalence, definitions and the relationship between the concepts are briefly presented before a detailed discussion in Sections 3.3.1-3.3.4.

The *tertium comparationis* is an assumed comparable common ground that exists between languages. It is an assumption that forms the foundation for a contrastive study which must have a common ground of reference across languages. Effectively, this ensures that a contrastive study compares like with like so that any findings or differences perceived can be interpreted faithfully. Equivalence is the criterion used to identify the most similar corresponding ways of expressing the same *tertium comparationis* in different languages. There is only one *tertium comparationis* in a study, but there can be many equivalences. The *tertium comparationis* is tested and analysed based on

equivalences, and the results of the testing of a *tertium comparationis* by an equivalence are the manifest similarities, also known as all the possible correspondences or ways of expressing an item in different languages. The goal is to use equivalences to test the *tertium comparationis* in order to measure these equivalences, reveal the manifest similarities, and choose the maximally similar correspondence.

To operationalise these concepts, the analysis presented in Chapter 5 reflects a good example. In Chapter 5, the contrastive analysis presented is based on the *tertium comparationis* or comparable common ground that questions are used as evaluative metadiscourse in the economics research article in English, French, and Spanish. This assumption involves finding all the questions used in each subcorpus in each language and testing them in terms of different equivalences that relate to question length, question function, and question location, for example – see Chapter 4, Section 4.3.2 for detail on the equivalences used to test the *tertium comparationis* in this study. In Sections 5.8-5.9, the manifest similarities and maximally similar correspondences of the analysis are presented.

While this brief example outlines the role and remit of these different concepts and terms, they are of course more complex; however, bearing this example in mind will be valuable in understanding the developments in the field discussed in the coming sections. To elaborate on these concepts, Section 3.3.1 elaborates on the concept of the *tertium comparationis* by outlining its role and value to contrastive studies as well as its evolution over time. Next, in Section 3.3.2, the concept of equivalence is discussed. The goal is to outline the types of equivalence used in contrastive studies and their role in testing the *tertium comparationis*. Finally, Section 3.3.3 discusses the challenges in measuring equivalences and outlines the concepts of manifest and maximum similarities.

3.3.1 The *tertium comparationis* in contrastive linguistics

The *tertium comparationis* has long been at the centre of debate in translation studies and contrastive linguistics alike. The *tertium comparationis* is the comparable common ground (Connor and Moreno 2005, p.156) and is the platform of assumed similarity from which a contrastivist or translator contrasts two or more languages. Contrastive studies might compare an item of language in one language with another item in another language on the assumption that they share a *tertium comparationis*. That is to say that the

assumption is that the two linguistic items behave similarly in some way *e.g.* they share semantic meaning. There can be only one *tertium comparationis* in a contrastive analysis, however the *tertium comparationis* extends to comparability at each level of analysis and comparison *i.e.* linguistic features share semantic meaning within specific and comparable contexts and texts. Moreover, this *tertium comparationis* cannot be tested in terms of itself as this could result in methodological circularity.

On this issue, Chesterman in his seminal work *Contrastive Functional Analysis* (1998) warns the contrastivist of putting the cart before the horse as he found that earlier contrastive approaches often suffered from methodological circularity (Chesterman 1998, p.53). The problem being, if two things are thought to be similar, the comparison could be designed to prove perceived similarity. Taking Krzeszowski's (1990, p.15) example of the rectangle, the issue of methodological circularity becomes clear. If the *tertium comparationis* in a comparison of a square and a rectangle were right angles, and the equivalence used to test that similarity was the presence of right angles, then they could be seen as the same, identical. Evidently, this is not true; however it can "proved" to be true using angles as equivalence and *tertium comparationis*. Conversely, were lengths and widths of the shapes considered as the equivalence to test the right angle *tertium comparationis*, the analysis would tell a different story. Acknowledging issues of such methodological circularity, Krzeszowski moved to try and solve this problem with his contrastive analysis approach.

Krzeszowski's (1990) approach to contrastive analysis was developed while contrastive analysis was moving from a formal to a functional view of language (Chesterman 1998, p.29). In such an approach, if the assumed *tertium comparationis* in comparing items of language were a formal or grammatical one, it is the job of the contrastive linguist to prove this *tertium comparationis* by identifying the extent to which the meaning or use is also the same (Krzeszowski 1990, p.19). This helps identify 'the closest approximations to grammatical word-for-word translations and their synonymous paraphrases' (Krzeszowski 1990, p.19). Chesterman, who built on this, established that the *tertium comparationis* is, in fact, not the start but the end point of a study (1998, p.52). In this way, an assumed *tertium comparationis* as he put it, would be vague and unfounded and could be proved through testing (1998, p.55). Others at the time, such as Goddard (1998)

and Wierzbicka (1999), turned to their work on natural semantic metalanguage as a means to stripping back a culturally laden *tertium comparationis* to universal semantic scripts (Wierzbicka 1999, p.238). For them, this can attest to the rigorousness of the *tertium comparationis* (Goddard 2012, p.711).

From a corpus-based contrastive analysis perspective, the *tertium comparationis* is what is comparable, and it has been argued that the *tertium comparationis* should be considered at all levels of the analysis to ensure comparability, and not just at the level of linguistic items being studied (Connor and Moreno 2005, p.161). For example, in studies such as Moreno (2004; 2008), Fløttum *et al.* (2006), McEnery and Xiao (2010), and Carter-Thomas and Chambers (2012), specialised multilingual corpora are used. In each case, the importance of maintaining comparability at each stratum of the corpus has been considered, producing an effective comparable common ground; for more on the relationship between corpus linguistics and contrastive linguistics in ensuring a rigorous *tertium comparationis*, see Section 3.5.4 for a detailed discussion.

In this study, the *tertium comparationis* is important as it has guided the thinking and interrogation of the data. In Chapter 2, Section 2.2, the comparability of academic English with other academic languages was problematised. So too was the comparability of the discipline of economics across languages in Section 2.3, and the comparability of genre of the research article across languages in Section 2.4. These are some of the strata of the corpora studied herein. The decision to focus on the economics disciplinary focus discourse community meant that the *tertium comparationis* was considered at all levels of analysis, including the wider socio-discursive context of the data. This ensured the comparability of one focus discourse community across English, French, and Spanish. Furthermore, at a textual level the *tertium comparationis* was further established in the decision to compare the economics research article across languages as the economics disciplinary focus discourse community's genre. Overall, the construction of the data presented in Section 4.2 outlines the comparability of the *tertium comparationis* in this study through a detailed description of the metadata at each stratum. However, for now what is important to note is that the theory underpinning the *tertium comparationis* has proved invaluable in choosing and designing the data analysed herein. Although, as Section 3.6 will show, this approach is not common among corpus-based contrastive

analysis studies where theories from contrastive linguistics are rarely employed. This study seeks to bring the fields of corpus and contrastive linguistics closer together in that regard, and the following section on equivalence shows how the *tertium comparationis* is tested in contrastive linguistics.

3.3.2 Theories of equivalence in contrastive linguistics

Moving on from the *tertium comparationis* towards a discussion of equivalence, it is worth noting that research within the field of contrastive analysis suffers from some terminological inconsistency, with words like congruence, equivalence, equivalent, and correspondence being used quite indiscriminately and polysemously within the literature. These terms are now briefly discussed and defined in terms of their use in this study.

Generally, equivalence is used to refer to the criterion used to test the *tertium comparationis* and to measure a degree of similarity or sameness in form and function depending on the type of equivalence used. However, for some others equivalence is purely for meaning and congruence purely for form (Kopczynski 1984, p.189). As well as this, correspondence for some has been used to discuss sentence level equivalence while equivalence is reserved for word level comparison (Fløttum *et al.* 2006, p.171; Johansson 2007, p.5). Some further confusion surrounds the identification of the result of an analysis as an equivalence (Adamska-Sałaciak 2010) or a correspondence (Wendland 2007, p.339).

For the sake of clarity, here, equivalences refer to criteria used to test *tertia comparationis* in order to identify correspondences as manifest similarities and choose a maximally similar correspondence. Initially, equivalence was concerned exclusively with comparing language as *langue*. This saw that equivalence was constrained to form and function where form is concerned with grammar and function is concerned with semantic meaning (Chesterman 1998, p.29). As contrastive linguistics evolved, so too did equivalence. James proposed equivalences such as translation equivalence, which relied on the ability of the contrastivist to identify sameness, a concept discussed in Section 3.3.4. He recognised an issue in identifying one true sameness (James 1980, p.169) and turned to Chomskian deep structure for answers. This gave James a foundation on which to found equivalences and correspondences.

Following this attempt to define equivalence, an important piece of research by Krzeszowski was presented in which he outlined his seven equivalences, namely: statistical equivalence; translation equivalence; system equivalence; semanto-syntactic equivalence; rule equivalence; substantive equivalence; and pragmatic equivalence (1990). Statistical equivalence is concerned with identifying the ‘maximally similar frequencies of occurrence’ (Krzeszowski 1990, p.27), whereas translation equivalence incorporates all kinds of translations based somewhat on what Krzeszowski terms translation performance (1990, p.161) or what translators do. For Krzeszowski, translation equivalence works parallel to semantic equivalence but need not be ‘constrained formally’ (1990, p.17), as in semantic equivalence with James (1980). He recognises that translation equivalence occurs when languages differ structurally or for stylistic purposes to which he attributes much breadth for analyses (1990, p.17). Next, system equivalence measures sameness paradigmatically in terms of word classes (1990, p.20); and semanto-syntactic equivalence, which builds on James above (1980), identifies sameness between structures on a deep structure level which is also formally and functionally constrained (Krzeszowski 1990, p.152). Following this, rule equivalence, which takes a transformation-generative grammar view, moves from deep structure to surface structure where surface structure is the representation of a sentence after deep structures have been transformed. From this perspective, rule equivalence identifies sameness at the phrase, clause, or sentence level in terms of congruence (Marton 1980) where congruent constructions are semanto-syntactically equivalent, contain the same number meaningful or lexical words⁷ which are systematically equivalent and syntactically equivalent (Krzeszowski 1990, p.135). Note that to achieve congruence in rule equivalence, Krzeszowski disregards the inclusion of closed-class words like auxiliary verbs or prepositions (1990, p.135). Next, substantive equivalence deals with sound and identifies sameness for phonological studies (Krzeszowski 1990, p.29); and finally, pragmatic equivalence, for Krzeszowski, identifies equivalence in terms of what is called maximum similarity (1990, p.218), a term which is discussed in Section 3.3.5.

Chesterman (1998), shortly after, presents his view of equivalence which is very much built on Krzeszowski’s work, despite some key disagreements. Chesterman, seeing the

⁷ Closed class words like auxiliary verbs are not counted (Chesterman 1998, p.34)

convergence of the fields of contrastive linguistics and translation studies, became increasingly confused by contrastivists' disregarding of translation as a simple performance (Krzeszowski 1990, p.161) when, for him it was clear that contrastive linguistics employed many translation strategies in its applications (Chesterman 1998, p.37). In fact, he identifies that many of Krzeszowski's equivalences mentioned above bare 'a striking resemblance to those recognised in translation theory' (Chesterman 1998, p.37) with the exception of rule equivalence. The table below illustrates this.

Table 3-1 Comparison of types of equivalence in contrastive linguistics and translation studies
(Chesterman 1998, p.38)

Krzeszowski's contrastive equivalences (1990)	Common translation equivalences (Chesterman 1998)
Statistical	Connotative, stylistic
Translation	Translation, textual
System	Formal
Semanto-syntactic	Formal and denotative
Rule	-
Substantial	Semantic denotative
Pragmatic	Dynamic, functional

He argues that although translating is indeed a performance, the ability to do so is a competence; this leads him to develop his theory of translation competence which requires 'the ability to move between two languages' (Chesterman 1998, p.38). His view of translation competence also leads him to move from generative views to functional views of language (Chesterman 1998, p.91) wherein a 'cross-linguistic [view of] equivalence is not absolute but a matter of judgement' (Altenberg and Granger 2002, p.10).

Beyond these key researchers, theories from different fields of study such as lexicography or linguistic typology propose further versions of equivalence. Many of these types of equivalence are comparable to those mentioned above, they simply differ terminologically and contextually. Adamska-Sałaciak, in her work on equivalence identifies that there are four different types of equivalence: cognitive, explanatory, translational, and functional (2010, p.397) which could be compared respectively to Krzeszowski's (1990) semanto-syntactic equivalence, translation equivalence, and pragmatic equivalence. Explanatory equivalence mirrors translation equivalence where it

tests *tertia comparationis* to identify correspondences that are structurally or lexically different but correspondent. Similarly, translational equivalence mirrors translation equivalence in the recognition of the need of the translator to use translation competence to identify corresponding items in different languages. This overlap occurring here is owing to the breadth in use that Krzeszowski lends his translation equivalence. Other types of equivalence exist such as prototypical equivalence (Gilquin 2008), which is comparable to semanto-syntactic equivalence (Krzeszowski 1990); or insertable equivalence (Karpinska 2015, p.79), which shares applications with translation equivalence (Krzeszowski 1990). This lack of uniformity in terms of terminology continues where, for example, Krzeszowski's rule equivalence (1990) is quite similar to Izquierdo's syntagmatic equivalence (2014), and Izquierdo's frequency equivalence (2014) bears a strong resemblance to Krzeszowski's statistical equivalence (1990). Bayar in her work on equivalence looked at formal equivalence, semantic equivalence, cultural equivalence, and pragmatic equivalence (2007). These types of equivalence echo those above with one distinction: cultural equivalence. This deals with the equating where possible of cultural features in different languages (Bayar 2007, p.177). However, this type of equivalence could be subsumed by Krzeszowski's translation equivalence or pragmatic equivalence, for example. Such overlaps occur throughout the interdisciplinary debate on equivalence rendering the term unnecessarily fuzzy. In an attempt to demystify equivalence, Table 3-2 builds on Chesterman's table above (1998, p.38) and has linked these various terms according to their usage based on Krzeszowski's seven equivalences where possible.

Table 3-2 Current types of equivalence in terms of Krzeszowski's seven types⁸

Krzeszowski's contrastive equivalences (1990)	Other types of equivalences with similar functions
Statistical	Connotative, stylistic, frequency
Translation	Translation, textual, insertable, explanatory, descriptive, cultural
System	Formal
Semanto-syntactic	Formal and denotative, lexico-semantic, cognitive, systemic, prototypical, conceptual, decontextualized, notional

⁸ Based on: James (1980); Zgusta and Hartmann (1984); Hausmann (1989); Krzeszowski (1990); Hausman and Werner (1991); Cop and Hausman (1991); Piotrowski (1994); Chesterman (1998); Wiegand (2005) and Adamska-Sałacia (2010)

Rule	No other equivalence identified
Substantial	Semantic denotative
Pragmatic	Dynamic, functional, lexico-pragmatic, functional, situational, communicative, discourse, cultural
Other	Lexicographic

Many of the researchers here have presented several types of equivalence which are important; not one, however, has done so as comprehensively as Krzeszowski (1990), whose approaches to contrastive studies remain important in contemporary research. For example, Izquierdo (2014) uses Krzeszowski's equivalences in a contrastive study measuring similarity and equivalence in an English and Spanish parallel corpus, ACTRES – research on the ACTRES corpus is discussed again shortly, in Section 3.6.

Overall, there has been a lot of controversy surrounding what constitutes equivalence as researchers have identified that their analyses are based, to a large degree, on bilingual (Krzeszowski 1990, p.161) or translation competence (Chesterman 1998, p.38). There has been a move away from a fixed or true view of equivalence to a view of equivalence as a perception, requiring the contrastive researcher to make an informed judgement (Chesterman 1998, p.16). Contrastive linguists make these judgments based on their language competence and expertise and for Chesterman, the role of translator competence is unavoidable and is something of a necessary evil (Chesterman 1998, p.16). Translator competence is a concept that is recognised in the literature (Fløttum *et al.* 2006) and as a limitation of this study which is discussed in Chapter 4, Section 4.6.2.

Overall, equivalence, like the *tertium comparationis*, is an important methodological concept in this study. Where the *tertium comparationis* is used to position and focus a contrastive analysis, equivalence is used to test the *tertium comparationis*. In this study, there are a number of equivalences used to test the use and function of questions as reader engagement which are detailed and linked to Krzeszowski's equivalences in Section 4.3.2. These equivalences are used to identify correspondences across languages which are identified through the measurement and comparison of equivalences.

3.3.3 Challenges in measuring equivalence in contrastive linguistics: towards maximum similarity

Issues of interpretation and measurement of equivalence are concerned with equating language pairs and sets in terms of what is called identity, where identity means ‘sameness’ (Stowe 2011, p.53). Many researchers problematised the idea of sameness, arguing that it is virtually impossible to say two words or phrases can fully share an identity in terms of form and meaning (Kopczynski 1984, p.285; Oleksy 1986, p.1416; Krzeszowski 1990, p.219; Neubert 1994, p.417; Chesterman 1998, p.24). For example, the meaning of the word “hot” in Ireland and *caliente* in Spain on a cultural level differ largely as what is hot for Irish people may not be hot for Spanish people, at least in terms of the weather. Therefore, equating words across languages can be challenging. In fact, arguably, it would be incorrect to identify synonyms within the same language variety as sharing an identity as they are not exactly the same (Adamska-Sałaciak 2013). For example, “big” and “large” in English are synonyms but they do not share identity as they are not the same. Furthermore, as the discussion of English, French, and Spanish academic language in Section 2.2.1 shows that the global position of these languages differs so greatly that to identify a sameness between them is virtually impossible. Evidently, sameness and identity are theoretical aspirations that would be difficult to put into practice. Therefore, it is unsurprising that research has moved away from this untenable idea of sameness (Adamska-Sałaciak 2013) and is now more concerned with identifying correspondences in manifest similarities and, following Oleksy (1986, p.1416), identifying the maximum similarity among the manifest similarities. Chesterman (2007a; 2007b) sees the importance of seeing manifest similarities and maximum similarity as the only means to assess a *tertium comparationis* across languages and this has led Chesterman to propose his view of convergent and divergent similarities (Chesterman 2007a, p.61).

The idea of convergence and divergence has been a part of equivalence studies for some time. In earlier studies in contrastive linguistics, researchers such as Fisiak termed this same distinction as theoretical and applied contrastive studies (1980). Theoretical is considered non-directional where studies look at the manifestation of universals in two languages, and applied is directional where studies are conducted in order to analyse a language in terms of another (Chesterman 1998, pp.40-41). Sovran’s identification of

convergence and divergence in similarity in 1992 built on this view and Chesterman (2007a; 2007b) arguably presents the most comprehensive overview of these two concepts.

Chesterman presents convergent similarity as the movement from two distinct items to one convergent one *e.g.* comparing *le/la/les* in French and “the” and the “zero article” in English in order to see if their meanings converge (2007a, p.61; 2007b, p.65). In contrast, divergent similarity begins with one item in one language and identifies correspondences in another language (2007a, p.61; 2007b, p.65). In each case there are limitations where in a convergent analysis, the *tertium comparationis* is the same linguistic feature in each language, for example questions. In a divergent analysis, the *tertium comparationis* could be the assumptions that questions in English have some correspondence in another language that may or may not be questions. In the case of each approach, equivalences are used to test the *tertium comparationis* in order to find items which behave similarly or differently. Taking this view, the identification of the maximum similarity is the next task which is concerned with finding a corresponding item of language that is as similar as possible to the item of language being studied. Importantly, this similarity is not constrained by similar forms.

For some like Kalisz (1981), identifying the maximum similarity is about identifying the degree and extent of similarity; however, for others it is about salient differences as opposed to degrees of differences (Krzeszowski 1990; Chesterman 2007a; Chesterman 2007b). Chesterman identifies that the several choices available to a contrastivist when contrasting two items in two languages are called manifest similarities (1998, p.68). To highlight this, the following example illustrates divergent manifest similarities available in English for the French phrase: *je comprends que*

I understand that

I appreciate that

I comprehend that

The manifest similarities have a similar meaning, but for subtle reasons each one is slightly different. This could be due to genre differences, register differences, regional differences, and differences over time, *etc.* The contrastivist must identify the most appropriate choice or the maximally similar correspondence which is guided by the mode,

register, varieties, genre, context, culture, and time in which the study is done. Manifest and maximum similarities are contextual, and therefore, similarity and difference within each equivalence is important, as a maximally similar correspondence could contain many formal linguistic differences, as the correspondence is not constrained to form.

Overall, identifying the maximally similar equivalence in manifest similarities necessitates choosing the most salient correspondence. In this view, as opposed to finding one equivalent word or expression that perfectly mirrors its counterpart, correspondence must be seen as a spectrum with various options and choices. This is reflected in more recent studies such as Rabadán (2005; 2011), Rabadán and Izquierdo (2013) and Izquierdo (2014) who look not to find correspondence but to measure different equivalences through testing the *tertium comparationis* and producing relevant manifest similarities. These studies are also important as they adopt approaches from both Krzeszowski and Chesterman, exemplifying their relevance to contemporary studies.

This study, like others such as Moreno (2004; 2008), Connor and Moreno (2005), Fløttum *et al.* (2006), Rabadán (2005; 2011), Rabadán and Izquierdo (2013), and Izquierdo (2014), recognises the important contributions researchers such as James (1980), Krzeszowski (1990), and Chesterman (1998; 2007a; 2007b), among the many others discussed here, have made to contrastive linguistics. In analysing questions in the economics research article in English, French, and Spanish, similarities and differences in each equivalence are discussed throughout Chapter 5 and Chapter 6. These findings help identify the manifest similarities for questions as reader engagement. Building on this discussion of the theoretical foundation of contrastive linguistics, the next section presents a review of methodological approaches in contrastive linguistics that have informed the methodological processes outlined in Chapter 4.

3.4 Methodological frameworks in contrastive linguistics

In recent and formalised methodologies in contrastive analysis, James (1980) has made noticeable contributions. His approach, although worthy of mentioning and in many ways pioneering, was limited and consisted of just two processes: description and comparison (1980, p.63). Beyond this, James identified the various types of contrastive studies that he considered possible *i.e.* lexical, grammatical, phonological, *etc.* (1980, p.61), and that each type of study would contain several sub-steps such as data assembling, data

description, data supplementation, and contrast formulation (1980, p.63). Unfortunately, his methodology did not account for the methodological circularity that Chesterman identifies (1998, p.53) which was discussed earlier in Section 3.3.2; however, it was an important first step in the development of a rigorous contrastive methodology following its earlier abandonment owing to theoretical weakness, as discussed in Section 3.2. It is important to note too, that although James drew the distinction between the *langue* and the *parole*, his view on *parole* was admittedly limited (1980, p.140). It was the beginning of a global shift towards the study of pragmatics and he imagined, quite rightly, that it would soon have an important impact on contrastive linguistics. He was right, and the next person to take this into account when developing a comprehensive methodology was Krzeszowski (1990).

Krzeszowski's work (1990) was and continues to be of great importance to contrastive linguistics. As discussed in Section 3.3.3, he was one of the first contrastivists to outline in detail the various kinds of equivalence available to the contrastivist and his approach remains important as it has continued to inform even contemporary analyses such as Izquierdo's work (2014). Following James' two-pronged approach, Krzeszowski added a middle step: juxtaposition – looking at languages side by side. The three-step approach describes the data that represents an assumed *tertium comparationis* (1990, p.35). The linguistic item in each language that forms the *tertium comparationis* is then juxtaposed in order to be tested by equivalences. Through the juxtaposition of the *tertium comparationis* to which other equivalences are applied, it is possible to measure equivalence and identify correspondences of the *tertium comparationis*. At this time, assessment of equivalence for both James and Krzeszowski relied on bilingual competences; however, for Krzeszowski, following Oleksy (1986), identifying equivalence moved from identity in James' case above (1980), to maximum similarity (Krzeszowski 1990, p.220) – concepts which were discussed in Section 3.3.4. However, Krzeszowski's approach was seen through a generative lens. This means that, like Marton (1980), deep and surface structures were important in identifying the maximum similarity. His generative view led to the development of Krzeszowski's extensive contrastive generative grammar theory (1990, p.169). In terms of methodological rigor, his approach did attempt to treat the *tertium comparationis* as assumed and unproven until other equivalences were applied, thus avoiding proving his assumption correct outright

(Krzeszowski 1990, p.19). However, his approach was based on the mutual constraints of form and function. Therefore, for him, testing form and function in terms of one another was sufficient (Chesterman 1998, p.30). This was the first attempt made in contrastive methodologies to avoid methodological circularity and was an important step forward in contrastive linguistics. Chesterman, who problematises contrastive methodologies at this time, commends Krzeszowski's attempt to attenuate this issue (1998, pp.30-31); however Chesterman does not consider Krzeszowski's effort sufficient (1998, p.52).

Chesterman, who admittedly builds on Krzeszowski's research (1998, p.40) went on to develop his contrastive functional analysis approach to contrastive analyses. There were two key aspects that raised this methodology to a higher level of rigor. The first, which has already been discussed in Section 3.3.4, was the transition from sameness to maximum similarity and the second was the importance he placed on falsifiability (Chesterman 1998, p.53).

For Chesterman, translation competence and subjective judgements (1998, p.39) are at the core of his approach. Translation competence is needed to identify a *tertium comparationis*, test it with equivalence, and prove a *tertium comparationis* through identifying maximally similar correspondences (1998, p.54). He proposed that contrastivists must put forwards a hypothesis that simply states the assumption that two things might be the same. Next this hypothesis must be falsified where the goal is not to prove the hypothesis right, but to prove it wrong by testing it with many types of equivalences (1998, p.54). Overall, contrastive functional analysis, which moved away from generative transformational grammar to a functional view of language, provided contrastive linguistics with a more rigorous methodology based on maximum similarity and falsifiability. The steps in a contrastive functional analysis are presented in Figure 3-1 (Chesterman 1998, p.54):

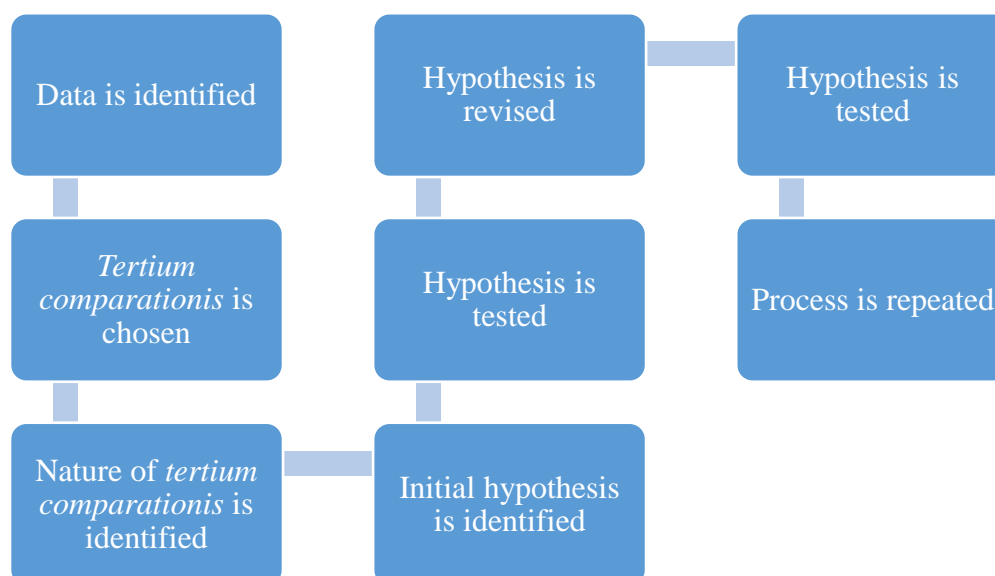


Figure 3-1 The process of a contrastive functional analysis

In contrastive functional analysis, there is only ever a hypothesis waiting to be falsified (Chesterman 1998, p.60) and this inevitably leads to weaker claims (Chesterman 1998, p.61); however, in search of a rigorous and comprehensive contrastive linguistic methodology, it is necessary. Contrastive functional analysis, in a sense, never truly ends and can be continuously tested as in this study, where two tests are applied to the same data in Chapter 5 and Chapter 6, refocusing and deepening the analysis of questions as reader engagement. Of course, more tests could be added; however, owing to limitations of space, there are only two analyses in this thesis.

This section on methodological frameworks in contrastive linguistics put into practical terms the theoretical evolutions detailed in Section 3.3. At the turn of the century, practices in contrastive linguistics had come a long way. Distinctions between *langue* and *parole* had been established. The *tertium comparationis* had been further theorised and equivalence had been transformed. Building on this, methodologies have been developed which have gifted the field of contrastive linguistics with greater authority and dependability than ever before. Bearing in mind the importance of contrasting language in contexts identified by James (1980), identifying *tertia comparationis* and testing and retesting them by Chesterman (1998), and testing them with a comprehensive set of equivalences like Krzeszowski's (1990), this study presents its methodological framework in Chapter 4 which borrows from all of these advances in the field. That being

said, the methodological framework applied herein does not only profit from advances in contrastive linguistics. With the advent of corpora and corpus-based contrastive analyses, work in contrastive linguistics has continued to develop, and the importance of the *tertium comparationis* and equivalences were further bolstered by the theoretical advances of corpus linguistics. As such, the following section reviews the field of corpus linguistics and outlines its relationship to contrastive analysis.

3.5 Corpus linguistics: background, theoretical foundation, and relevance to contrastive linguistics

Corpus linguistics is a research approach that has come to support ‘empirical investigations of language [...] use’ (Biber and Reppen 2012, p.ix) through the analysis of large collections of texts called corpora. Corpora have long been used to analyse and identify patterns in a language or a genre (Biber *et al.* 1999, p.4), although they have not always been so called (McEnery and Wilson 2001, pp.2-3). Long before the advent of computers, empirical language data used in contrastive research like those of Fries (1945) or Lado (1957) mentioned in Section 3.2, could be considered corpus data. In fact, any research in linguistics or applied linguistics that has used collections of empirical language data could be said to have used corpus data to some extent. Nowadays, corpora are machine readable or digital and corpus linguistics is underpinned by theoretical concepts and methodological processes that have advanced and informed the field.

This section discusses the field of corpus linguistics by first outlining its historical context and the current state of the art in Section 3.5.1. In so doing, parallels are drawn between the advances in corpus linguistics and contrastive linguistics in order to demonstrate how the current state of the art of corpus linguistics complements contrastive linguistic approaches. Following that, Section 3.5.2 discusses the core concepts that underpin corpus linguistics and the types of corpora that exist within the field. This serves to outline concepts like representativeness and discuss literature evaluating corpora and the types of corpora that exist. Following that, the methodological processes used within corpus linguistics are discussed in Section 3.5.3, focussing in greater detail on those processes most relevant to this research. Finally in Section 3.5.4, multilingual corpora and the relationship between corpus linguistics and contrastive linguistics are discussed.

3.5.1 Corpus linguistics background and current state of the art

In the 1950s, the value of early corpus data, which was widely used though would not resemble what we see as corpus data today, was called into question in the dispute between empiricists and rationalists. Empiricists are concerned with real language use as data (Sampson 2013) and see corpus data as important for the development of grammars, dictionaries, and teaching resources, for example (Kennedy 2014, p.47). When seeking examples of language use, empiricists draw from documented language use as opposed to inauthentic language prescribed by one or few rationalist thinkers (McEnery and Wilson 2001, p.5). Rationalists however, base their decisions not on empirical data but on the development of ‘a theory of mind’ (McEnery and Wilson 2001, p.5) and an ‘innate dispositional knowledge of language’ (Smith and Allott 2016, p.218). They conceptualise and theorise rules and patterns in language, test these theories and make claims based on their own understanding of language.

The rationalists’ arguments, invigorated by Chomsky (Smith and Allott 2016, p.219), carried weight as rationalists considered corpus data as a poor representation of what Chomsky called language performance (1964). As mentioned in Section 3.2, in a Chomskian view, language could be seen in two ways, as competence and performance (1964, p.35). Competence refers to our deep understanding of the complexities of language and performance refers to the actual language produced (Chomsky 1964, p.36). For Chomsky, language competence is the end goal for a language learner, whereas performance is merely its shadow. He criticised corpus data, arguing that data comprising entirely of performance language could never endeavour to help a learner unlock their language competence (McEnery and Wilson 2001, p.12). He also argued that corpus data is finite and could not account for all instances of language since corpora are, in their nature, incomplete. For Chomsky, a large collection of incomplete performance language carried no real merit (McEnery and Wilson 2001, p.12). Instead, he, as a rationalist, argued that linguists should avoid using such data, and instead should attempt to theorise language based on their own language competence and innate understanding of language and linguistics (Chomsky 2009, p.97). Therefore, owing to the conceptual difficulties with corpus data and the lack of standardisation surrounding the collection and application of corpus data, the role of corpus data in language research diminished, with little research being conducted on corpora until the 1980s.

Quite distinctly, corpus linguistics has been growing expansively since the 1980s, most probably due to technological advancements in the field (Altenberg and Granger 2002, p.3) *i.e.* the wider availability of technology. This growth has seen the application of corpus linguistics to myriad applied linguistic fields of study and McCarthy and O’Keeffe note that corpus linguistics has implications ‘in [the] pursuit [of research] in areas such as language teaching and learning, discourse analysis, literary stylistics, forensic linguistics, [and] pragmatics’ (2010, p.7). To this list, contrastive linguistics could be added (Granger 2003, p.18) which saw a renaissance in the 1990s in the form of corpus-based contrastive analysis (Marzo *et al.* 2012, p.1). Reflecting back on the discussion in Section 3.2 of this chapter, the evolution of corpus linguistics mirrors that of contrastive linguistics in many ways. As contrastive linguistics moved from *langue* to *parole* and abandoned Chomskian views of competence and performance, it found itself in a similar position to corpus linguistics that valued language in context. Like contrastive linguistics, corpus linguistics was criticised for being theoretically weak and the movement sparked the need for further theorisation. Of course, technological advances supported this greatly; however, computers were not enough to overcome the issues that surrounded the quality of corpus data. Corpus linguistics needed to revolutionise and in so doing address the rationalist grievances. This was achieved through taking a more refined perspective on what constitutes a corpus in terms of methodological concepts such as size, representativeness, and sampling. These concepts as well as the types of corpora they help create are discussed in the following section.

3.5.2 Corpora in corpus linguistics: theoretical underpinning and types of corpora

Addressing the rationalistic concerns above, the movement away from language as *langue* and towards language as *parole* repositioned corpus linguistics as a valuable field in the study of language in context. However, the finite nature of a corpus remained problematic, owing to its limited size and the language it represents. However, this issue becomes less problematic when the concepts of representativeness and sampling in corpus construction are considered (Biber 2012, p.30). A corpus is representative of a particular population depending on the sample collected (McEnery and Wilson 2001, p.30). This is achieved, as in the case of this study, by choosing data of a shared discourse community, genre, time-frame, sampling period and data with minimal outliers in terms of content and demographic information. For example, in the case of this study of

economics research articles in English, French, and Spanish which were published between the years of 1995 and 2003, the results cannot tell us anything about academic writing in general nor can they tell us about academic writing in another discipline, language, or time. The corpus is representative of this population and nothing more (Biber 2012, p.4) and the strength of representativeness resides in the accurate and detailed documentation of corpus metadata.

Considerations of representativeness very much reflect the challenges faced in contrastive linguistic studies surrounding the identification of the *tertium comparationis*. Representativeness is concerned with capturing metadata at all levels of the corpus in order to define the population. When comparing across languages, *tertium comparationis* must be considered at all levels of the data to ensure comparability *i.e.* corpora in different languages must reflect the same metadata and therefore represent comparable populations (Connor and Moreno 2005, p.161).

Representativeness in corpus linguistics is a complex concept, and very much reflects the complexities that exist in data collection within the wider data sciences (Shah *et al.* 2015, p.9). That is to say, there are not only difficulties in deciding what data to capture but challenges in evaluating the rigorousness of data collection and description. This PhD thesis recognises the continued evolution within the field of corpus linguistics and data sciences. This brings with it challenges surrounding defining what is representative. This study profits from the current state of the art of corpus linguistics and its application to contrastive linguistics by interrogating the quality of the data (Taylor and Marchi 2018). Overall, building on the concepts of representativeness and sampling, a taxonomy of corpus types has emerged in the literature. These types of corpora are now briefly discussed.

Generally speaking, corpora can be representative of larger balanced language sets or extremely specific selections of language. Furthermore, upon deeper classification, corpora tend to capture language of specific communities, expert speakers, multilingual communities, and learner and non-expert users of language. All corpora are so defined based on the concept of representativeness which dictates whether a corpus is considered general or specialised based on the data it captures (McEnery and Xiao 2008, p.21; Cheng 2012, p.166).

Learner corpora for example can be important in understanding how learners of different languages use language, and although learner corpus research is a relatively new field, it has grown quickly and substantially. Studies of learner corpora identify the specific needs of learner populations as they contain ‘a systematic computerized collections of texts produced by learners’ (Nesselhauf 2006, p.128), such as ICLE – the International Corpus of Learner English (Granger *et al.* 2002). Research in this area views learner language as a ‘dynamic and independent language system’ (Chau 2012, p.194) and has focused directly on how learners write, for example, as in Liardét’s study of high performing learners’ uses of grammatical metaphor and how it differs from that of low performing students (2016) or indirectly as in McCarthy’s work on the role of learner corpora in informing language teaching resources (2016). Such research has been extremely important in language teaching and learning – for more detail on learner corpora and their role in informing language teaching, see Nesselhauf (2006), Granger (2009; 2012), Chau (2012), Pérez-Paredes and Díez-Bedmar (2012), Frankenburg-Garcia *et al.* (2013) and O’Keeffe and Mark (2017).

General corpora, such as the BNC⁹, the Spoken BNC2014¹⁰, or COCA¹¹ are usually quite big and contain many sub-genres metadata, and a greater variety of samples of language (McEnery and Xiao 2008, p.21). Specialised corpora like the KIAP corpus, however, are typically much smaller and are often composed of specific ‘registers, genres [or] language forms’ (Cheng 2012, p.166) such as academic writing (Fløttum *et al.* 2006). Specialised corpora have come to have extremely useful pedagogical applications (Rodgers *et al.* 2011, p.46) and have at times been referred to as pedagogic corpora (Willis 1998).

As they represent a specific population, and more specifically, the language of a specific population (Gavioli 2005, p.55), specialised corpora can serve to expose the features and language of a specific language community (Rodgers *et al.* 2011, p.394). For example, Li and Qian (2010) look at the use of Coxhead’s academic wordlist (2000) in teaching financial language or Poole’s work on corpus-informed teaching of rhetoric for academic purposes (2016). These studies on specialised corpora can be used to inform very specific language courses *e.g.* a language course for medical purposes (Lu and Corbett 2012),

⁹ The British National Corpus, (BNC 2007).

¹⁰ The British National Corpus 2014 (Love et al. 2017).

¹¹ The Corpus of Contemporary American English (Davies 2010).

language for business purposes (Tangpijaikul 2014), or language for academic purposes (Flowerdew and Forest 2015). Moreover, the use of specialised corpora can help identify the tendencies and linguistic practices within different discourse communities in terms of varying metadata like member type, text type, time-frame *etc.*

Learner corpora, general corpora, specialised corpora, and multilingual corpora (which are discussed in Section 3.5.4) each serve different purposes and hold different value in the field of corpus linguistics. What aligns them is their theoretical underpinning in terms of their principled construction as representative corpora. An important consideration of corpus data is that comparability across corpora is less often considered than in contrastive linguistics. This is because comparing databases that are not necessarily comparable is not uncharacteristic of corpus linguistics (Lindquist and Levin 2000). For this study, comparability is important, therefore, margining these two fields is key; this is discussed in Section 3.6. Overall, the quality of a corpus is determined by the underpinning theory of community and how the reflects the population and metadata captured. Corpora, no matter the type, are analysed using corpus analysis software such as AntConc (Anthony 2018) or Sketch Engine (Sketch Engine 2019). These software allow for a number of different corpus analysis functionalities which are discussed in the following section.

3.5.3 Methodological processes in corpus linguistics

In conducting a corpus-based contrastive analysis of academic writing in English, French, and Spanish economics research articles, this research profits from several applications of corpus analysis software. Many researchers have discussed the processes in corpus linguistics that can be used to analyse language and there is little controversy surrounding their functions, perhaps only to the extent to which they are used and their application to and interpretation of different data sets (O'Keeffe *et al.* 2007; Baker 2012; McEnery and Hardie 2012; Walsh 2013). In any case, the processes are computed through a range of corpus analysis software.

Corpus analysis software/interfaces such as Sketch Engine (Sketch Engine 2019), #LancsBox (Brezina *et al.* 2019), and AntConc (Anthony 2019) are applications used by corpus linguists to analyse corpora. They contain functionalities which allow for the study of keyword lists, clusters and n-grams, lexico-grammatical profiles, concordance lines,

and frequency and word lists, for example (O'Keeffe *et al.* 2007, pp.8-16; Baker 2012, p.4; Walsh 2013, p.45). These software vary by degree of complexity and accessibility. Sketch Engine, for example, requires a subscription for use and allows for complex processes. In its conception, Sketch Engine was designed for lexicography (Kilgarriff *et al.* 2014), and it has auto-tagging software for multiple languages such as the Robust Accurate Statistical Parser (Briscoe *et al.* 2006), using the CLAWS2 tag set. This allows for effective parts of speech searching within the software. Sketch Engine also contains a number of automatic statistical computations, powerful list and filter functionalities, space for detailed metadata descriptions, and unique features for the analysis of collocations and word sketches. Thomas (2015) details the many functionalities of Sketch Engine. Other software like #LancsBox (Brezina *et al.* 2019) work similarly and offer unique data visualisations and built-in specialised linguistic statistical packages produced by Lancaster Stats Tools Online and detailed in Brezina (2018).

Despite their evident value for corpus linguistic studies, neither Sketch Engine nor #LancsBox were used herein for a number of reasons. While the main focus of this study is the linguistic comparison of questions across languages and the production of a linguistic description, it is hoped that the linguistic insights produced will be of use in educational contexts. Therefore, Sketch Engine which requires subscription and is for extremely specialised users would make replication of this study by pedagogical specialists challenging. Moreover, at the time of conducting this project and analysing the data, #LancsBox was not fully developed and eligible for use. Therefore, this study used the software AntConc (Anthony 2019), which is freely available online, easy to use, and comes with a range of products that work with the software. This affords AntConc a reflexive complexity where it can compute more complex processes or remain more accessible for non-experts depending on the task. Each one of these software computes statistical information on corpora and their statistical packages differ. This reflects the current debate in corpus linguistics surrounding the choices of statistical analyses by corpus linguists which remains highly contentious (Brezina 2018; Gabrielatos 2018; Gries 2019). Given that the corpus used herein is relatively small, this study has largely excluded complex statistical analyses which are more effective and accurate on larger data sets (Brezina 2018, pp.20-21). Overall, AntConc allows for the computation of the core processes used in corpus linguistics to analyse language (O'Keeffe *et al.* 2007; Baker

2012; McEnery and Hardie 2012; Walsh 2013) and for the purposes of this study, is sufficiently sophisticated. However, a caveat to using AntConc is that it is necessary to capture metadata in accompanying documents like Microsoft Excel spreadsheets as the software is not powerful enough to process and search within metadata. These core methodological processes of corpus linguistics are now discussed.

Frequency lists are the most used tool in corpus linguistic research (Gries 2012, p.111). They show the most and least frequent words in a corpus which is indicative of their frequency of use within the relevant population or sample (Tognini-Bonelli 2001, p.4). Depending on the software used, lists are filterable and searchable to identify lemma forms, rank and frequency, and, as with Sketch Engine, can be used to produce a frequency list of a set of searched words only. These frequency/word lists are extremely useful and have informed many studies such as Farr's study of response tokens in student-tutor meetings (2003) which used frequency to determine the main response tokens identified. Of course these frequency lists are limited to counting forms and therefore cannot count forms constrained by function without extra metadata tagging. Beyond identifying lists of words, cluster analyses and the study of n-grams involve the analysis of words and their role in larger chunks of language through a multivariate statistical analysis (Friginal *et al.* 2014, p.1) where multivariate means that there are several variables in a study. Corpus analysis software can compute long multi-word combinations which give a better understanding of how words work together to create meaning (O'Keeffe *et al.* 2007, p.13; Holub *et al.* 2012; Gilquin 2015).

While frequency lists and n-grams identify language items in isolation, corpus linguistics is also concerned with language in context. Concordance lines or keyword in context (KWIC) searches, which are arguably the next most common way to represent corpus information (Cheng 2012, p.6), allow for this. KWIC searches have been extremely useful for 'analysing word meaning in a corpus' (Deignan 2012, p.16). KWIC¹² searches take the searched word or phrase, better known as the node, and present each occurrence of the node centred on the screen with words coming before and after. These concordance lines help researchers identify the functions of the language being analysed, as evidenced

¹² Note that here KWIC, like for McEnery and Hardie (2012, p.35) can refer to both a word and a chunk of language.

in studies such as Chambers' work on the language of journalistic discourse and the French word *jouer* (2011). It is important to note that although KWIC searches can present many useful examples of language patterns, this would be meaningless without effective and rigorous interpretation by the researcher in a horizontal-reading methodology which considers the language in its wider context (Aijmer and Rühlemann 2015, p.9).

Concordance lines can also be used to create what O'Keefe *et al.* call lexico-grammatical profiles (2007, p.14). These profiles serve to analyse a node according to 6 facets: its collocations *i.e.* words that frequently occur together; its role in chunks or idioms; its syntactic restrictions *i.e.* sentence position or potential collocates; its semantic restrictions *i.e.* words it is never used with; its prosody *i.e.* ability to imbue a sentence, a chunk or itself with positive or negative connotations; and finally, any other relevant or interesting facets of the word. These lexico-grammatical profiles are a useful means to analyse language in corpora which can serve to better inform grammars and dictionaries. Studies that undertake elements of lexico-grammatical profiling have analysed collocation and colligation and their representation in corpora and learner dictionaries (Wojciechowska 2016), the syntactic values of twitter hash tags (Rueda and Alamán 2015), or changes in semantic prosody in diachronic studies of Persian (Hosseini-Maasoum *et al.* 2016). All of these studies of concordance lines rely on effective sampling, filtering, and sorting functionalities.

Other functionalities like keyword analyses are one of the most common significance tests employed in corpus linguistics (McEnery and Hardie 2012, p.51). This significance test is conducted through comparing word lists from the target corpus with one from a larger reference corpus (O'Keefe *et al.* 2007, p.13). In doing this, the significance of each word in the corpus is tested in order to identify which words could be considered key (McEnery and Hardie 2012, p.51); however, again this is constrained by significant form and not function unless tagged explicitly. Key words are not the most frequent words but have been considered to be the most unusually frequent words in a corpus (O'Keefe *et al.* 2007, p.13). Others such as Evison move to further develop our view of keyness through identifying both positive and negative keyness (2010, pp.127-128). For her, positive keyness captures words that are unusually frequent in the target corpus whereas

negative keyness refers to words that are unusually infrequent in comparison to their frequency of use in the reference corpus (Evison 2010, pp. 127-128). Given that this study compares multilingual corpora and not a corpus with a reference corpus, and is concerned with function as well as form, keyness as a measure has not been implemented here. Though future studies could look to English, French, and Spanish reference corpora to further triangulate this study.

Of these core processes, this study makes use primarily of frequency and cluster analyses, as well as concordance searches. However, it must be noted that these searches make use of technical tools that are constructed with a philosophical perspective on the relationship between statistical measurements and language phenomena. Corpus linguistics has been rightly criticised for unquestioningly using these software and for accepting statistical information they deliver without thorough consideration (Gries 2010, p.269; Brezina 2018, p. 283). Therefore, alongside the KIAP-EEFS corpus, detailed metadata have been captured to guide in the interpretation of its data. Moreover, theories of comparability from contrastive linguistics have allowed this thesis to attest to the quality of the data. These data, metadata, and the methodological processes used in this study are outlined in detail in Chapter 4. These processes are applied to the KIAP-EEFS corpus, which is a multilingual corpus. Therefore, the following section discusses multilingual corpora and the relationship between corpus linguistics and contrastive linguistics.

3.5.4 Corpus linguistics and contrastive linguistics: multilingual corpora and complementary foundations

Corpus-based contrastive analysis, as the name implies, merges the fields of corpus and contrastive linguistics, and research in this area profits primarily from analysing either learner corpora or multilingual corpora. This study is concerned with multilingual corpora, of which there are three main types: (A) source texts and their translations; (B) monolingual subcorpora designed in different languages using the same sampling frame and (C) a combination of these two types (McEnery and Xiao 2008, p.18). Following McEnery and Xiao's (2008) terminology, (A) is as parallel corpus and (B) is a comparable corpus (2008, p.19), and McEnery and Xiao also refer to (C) as a comparable corpus (2008, p.20), which risks causing confusion. However, it is not important here as this study uses only type (B) which henceforth shall be referred to as a comparable corpus.

Comparable corpora are multilingual corpora that do not contain translations but contain texts ‘collected using the same sampling frame’ (McEnery and Xiao 2008, p.20). The texts are equal in proportion according to language and genre and are collected within ‘the same sampling period’ (McEnery and Xiao 2008, p.20). These corpora necessitate careful construction *i.e.* the framework for their construction must be rigorous and consistent if any results drawn from them are to be considered accurate or dependable (Altenberg and Granger 2002, pp.8-9; McEnery and Xiao 2008, p.21). Upon successful construction, these corpora are seen to be representative of varieties of languages or genres (Biber 2012, p.1). In Section 3.5.2, sampling and representativeness are presented as key characteristics of well-constructed corpora. Therefore, comparable corpora must also be representative of a specified language sample (Altenberg and Granger 2002, pp.8-9; McEnery and Xiao 2008, p.21). Moreover, given the importance of the *tertium comparationis* to the data being compared across languages, comparable corpora must be comparable at each stratum of the data (Granger 2010, p.5). Connor and Moreno identify the importance of ensuring a *tertium comparationis* at the following strata of a comparable corpus (2005, p.161):

- | | |
|---|--|
| 1. Text form | 11. Predominant text types |
| 2. Genre | 12. Format features: |
| 3. Mode | <ul style="list-style-type: none"> • Length • Intertextuality • Visual features |
| 4. Participants: | 13. Global communicative event |
| <ul style="list-style-type: none"> • Writer/speaker • Audience/target | 14. Setting |
| 5. Situational and dialectal varieties | 15. Purpose of communication |
| 6. Tone | 16. Global rhetorical strategy |
| 7. Channel | 17. Overall subject matter of the |
| 8. Point of view | topic, its discipline |
| 9. Level of expertise | 18. Global superstructure |
| 10. Textual units of analysis | |

Comparable corpora are often specialised (Eckart and Quasthoff 2013, p.152). Though generally small in stature, specialised corpora can boast advantages and it has been argued that small genre-specific corpora can be heavily patterned and can thus produce valuable insights into language use in the respective areas they represent (Aston 2001, p.30). Therefore, challenges in faithfully comparing linguistic items and features across languages, discussed throughout Chapter 2 and Chapter 3 so far, can be attenuated

through the construction of specialised comparable corpora (Altenberg and Granger 2002, p.8). The importance of interrogating the data in terms of what they represent is key which is why comparability across languages was an important consideration in Chapter 2, Section 2.2-2.4. Overall, comparable corpora are a rich applied linguistic resource and their applications are manifold. Some areas that have largely invoked the use of specialised comparable corpora are language teaching and language for specific purposes (Van Bonn and Swales 2007) as well as contrastive linguistic descriptions like this thesis (Aijmer and Lewis 2017).

Overall, in Section 3.5.1 it became clear that by reconceptualising the field and making advances in technology, corpus linguistics could make new offerings to fields in linguistics. In this section, it has been made clear that multilingual and comparable corpora in corpus and contrastive linguistics hold a common ground and these fields share an importance placed on the conceptualisation of comparable and representative data. This has resulted in an inter-disciplinary field of study that has come to be known as corpus-based contrastive linguistics (Marzo *et al.* 2012, p.1). Current trends in contrastive linguistics notably profit from analysing corpora across language pairs or sets and so too does this thesis. The following section moves on from corpus linguistics to a discussion of the current state of corpus-based contrastive analysis.

3.6 Corpus-based contrastive analysis, academic discourse, and the current state of the art

As presented in Chapter 2 Section 2.2, most studies on academic writing have analysed the English language (Flowerdew 2002, p.2) and in the case of contrastive linguistics, most corpus-based contrastive analyses have been conducted on English and another language (Marzo *et al.* 2012, p.1). This is likely owing to English's 'privileged position [...] in the world of science today' (Fløttum *et al.* 2006, p.25). English's role as the lingua franca of academia has encouraged researchers to accept the view that comparing English academic writing with that of other languages can produce useful information for those who wish to publish in English (Mur-Dueñas 2011, p.3068; Shi 2002, p.628). Unlike monolingual studies, contrastive studies in academic writing form part of a much 'more modest' (Carter-Thomas and Chambers 2012, p.17) collection of studies due to the fact that research in contemporary corpus-based contrastive linguistics is a relatively recent

endeavour. Nonetheless, in a world of ‘increasing international exchange’ (McEnery and Xiao 2008, p.18) where ‘interlingual and intercultural communication’ (Granger 2003, p.18) is of growing importance, corpus-based contrastive analyses of academic discourse across various language pairs are becoming increasingly common. It is in this field of corpus-based contrastive analysis that this study is situated and to this field that this study contributes.

The growth of this field does not come without criticisms, where contrastive studies have been criticised for issues surrounding comparability in terms of sample size, sample genres or unclear *tertia comparationis* (Hyland and Salager-Meyer 2008, pp.320-321). The research group ACTRES – Análisis Contrastivo y Traducción English-Spanish – which ‘has dealt with an extensive array of cross linguistic phenomena’ (Izquierdo 2014 p.141), has presented interesting and comprehensive approaches to corpus-based contrastive analyses, effectively merging the two fields of contrastive and corpus linguistics. Rabadán (2005; 2011), Rabadán and Izquierdo (2013), and Izquierdo (2014) have produced corpus studies that have built upon Krzeszowski (1990) and Chesterman’s (1998) methodologies, discussed in Section 3.4, demonstrating their incorporation into corpus-based contrastive linguistic approaches.

ACTRES centre their philosophy on modern views of similarity as a rejection of identity and the adoption of convergent and divergent approaches, and manifest and maximum similarities. This very much draws from Chesterman’s school of thought (2007a). Starting with a convergent similarity as *tertium comparationis*, for example, they move to take two items of language in two languages and try to find correspondences (Izquierdo 2014, p.146). This convergent similarity triggers the need for a corpus-based contrastive analysis of a comparable corpus. Their inductive approach (Izquierdo 2014, p.144) centres on a functional approach to language through an analysis based on Krzeszowski’s equivalences (Rabadán and Izquierdo 2013, p.48; Izquierdo 2014, p.141), thus hybridising Krzeszowski and Chesterman’s approaches to contrastive analysis. To the Krzeszowskian-style equivalences, they apply a Chesterman-style contrastive functional analysis framework based on testing and falsifiability (Rabadán and Izquierdo 2013, pp.48-49). Starting with their hypothesis and *tertium comparationis*, in their case based on function, they wish to measure equivalences. They analyse their data in terms of

syntagmatic, paradigmatic, and frequency equivalences *i.e.* Krzeszowski's rule, system, and statistical equivalences respectively. They apply these equivalences through Chesterman's contrastive functional analysis, testing and then measuring correspondence using their own tool which measures the degree of functional similarity based on usage. All of this contrastive work happens within a corpus linguistic paradigm where the construction of the data analysed and the processes used to analyse the data are guided by corpus linguistic approaches. While their work builds on and adds to contemporary approaches to contrastive analysis, their equivalences are more specific and limited than Chesterman's or Krzeszowski's, and as such, are not employed herein. Moreover, their framework limits to function as a *tertium comparationis* which is not the case in this study. Nonetheless, their work is among the most comprehensive in methodological descriptions that draws from both contrastive and corpus linguistics.

Within the field of academic discourse, the following section, Section 3.6.1, presents a review of published research that take a corpus-based contrastive analysis approach and contain one or more of the languages studied in this PhD. Owing to the sparsity of contrastive research on the economics research articles in English, French, and Spanish, the scope has been widened to a review of studies that focus on academic discourse in general, and not only the discipline of economics and the research article genre exclusively. Following that, Section 3.6.2 identifies the gaps in the field of corpus-based contrastive analysis, highlighting the contribution of this study and the apparent future directions of the field.

3.6.1 Corpus-based contrastive studies of academic discourse: a review

As this PhD is based on a contrastive analysis of questions in economics research articles in English, French, and Spanish, it has drawn from relevant contrastive research on these languages in the context of academic discourse. The purpose of this section is to present a review of the collection of corpus-based contrastive analyses that have informed this study, signal relevant findings and methodological considerations, and highlight the body of work to which this thesis contributes. Given the sparsity of research on languages other than English, some of the research studies presented here are the same as those discussed in Chapter 2, Sections 2.5-2.9. However, the focus in this section is not concerned solely with the results but on the relative description of languages in terms of one another.

When considering studies on English and French, differences in the use of first person pronouns have been identified in studies on KIAP (Rongen Breivega *et al.* 2002) where notably fewer first person pronouns are used in French than in English or Norwegian. This presents an argument that self-mentions are more characteristic of English than French. This positions English as interpersonal while French appears to show little interpersonality. Similar findings emerge in the comparison of academic vocabulary in English and French where French appears to not have as varied a lexis in academic writing (Cobb and Horst 2005). Others such as Fløttum *et al.* have conducted divergent analyses, where their study of “let us” in English, French, and Norwegian is based on quantitative analyses involving complex non-parametric statistical methods and qualitative analysis which focuses on English as a point of reference for the interpretation of “let us” imperatives and their formal correspondences in French and Norwegian (2006, p.169). In categorising these “let us” imperatives according to eight functions across the three languages by considering their pragmatic equivalence, they then consider their roles in the text and explanations in order to identify any differences. Overall, they argue that English is more explicit in nature than French.

Further multilingual studies have considered differences across English, French and German and their research indicates important difference in linguacultures *i.e.* language and academic registers differ at the cultural level (Siepmann 2006). This focus on academic cultures across multiple languages does not position languages in a binary contrast, but as relative to one another, which allows for the identification of more nuanced cultural practices. In terms of evaluation in French and Italian linguistics research articles, evaluative markers may sometimes take the form of one or multiple words which appears to be a common practice across romance languages (Rentel 2006). In English and French abstracts, there appears to be a number of important differences between the two languages. For example, there are notable differences in the use of voice, pronouns, sentence length, and reporting (Van Bon and Swales 2007), and there is evidence of more passives in English than French, that English is less critical than French, and that English sentences are shorter than French sentences (Van Bon and Swales 2007). Tutin (2010b) finds that conclusions in the English academic research article in economics make greater use of evaluative adjectives than French, and Grossmann and Tutin (2010b) find intertextual features of writing in the economics research article are

more common in English than French when referring to graphs and tables. These studies serve to effectively show how these languages compare to one another, but it is difficult to ascertain from these studies how evaluation English and French is situated more broadly across multiple languages. In a similar note, Salager-Meyer (2011a) argues that research papers in French are more prescriptive and authoritarian than English and that research papers in English tone down their claims to avoid face threatening acts. Others have identified that French uses fewer first person pronouns in the introduction than English (Carter-Thomas and Chambers 2012) and Curry and Chambers' (2017) study of questions in linguistics research articles reveals that questions are more common in French than English, and that French uses more questions when framing the discourse and asking real questions, while questions used to organise the text and set up claims are more common in English. They also find that French uses more content and polar questions and makes more varied use of forms with content, polar, and declarative questions occurring more frequently. What these studies show is a clear relationship between English and French that views them as a binary. Largely, English is seen as interpersonal and French is seen not to be interpersonal, exceptions notwithstanding. However, this binary view is problematic and does not account for linguistic relativism, where most likely interpersonality is better understood as a spectrum as this thesis posits in Chapter 7, Section 7.2.9.

Focussing on English and Spanish, research has shown that there is more metatext in English economics research articles than in Spanish (Valero-Garcés 1996). In terms of passivity, research has shown that the *se*-passive in Spanish is more likely to correspond to the English passive voice in academic writing and that passive simple tenses function similarly, in terms of creating distance and impersonality, in both languages (Espinoza 1997). Looking at the register of English and Spanish report writing, the incorporation of metaphor has been noted and its use varies across languages where Spanish prefers the use of metaphors of personality while English sees more nautical metaphors like 'weather the storm' (Charteris-Black and Ennis 2001, p.257). In Spanish abstracts in social sciences, there is evidence that Spanish largely follows the international and English-based conventions (Martín 2003) where most research articles in Spanish mirror the English language structure of IMRD, as already discussed in Section 2.7.1. There is also comparative evidence of passivity being more common in methodology sections in

English than Spanish. However, this does not prove that Spanish does not use passivity at all, or that Spanish uses it the least. This is because the comparison of two languages makes it difficult to understand these findings in a wider context. Others like Moreno (2004; 2008) and Connor and Moreno (2005) analyse comparable corpora, investigating cross-cultural differences and L1 interference in L2 writing in English and Spanish. For them, cross-cultural differences can be explained to some degree through their classification as ‘values, norms, learning processes and educational trends’ (Moreno 2008, p.26) within specific cultures. Their work places importance on comparability and ensures the *tertium comparationis* is considered at each level of corpus construction in terms of text form, genre, mode, population, format features, global communicative event, setting, general purpose of communication, global rhetorical strategy, overall subject matter of the topic and discipline, and textual units of analysis and predominant text types, among others (Connor and Moreno 2005, p.161). This is reflected in Moreno (2004) where English and Spanish research articles are analysed in terms of the language used by authors when concluding. She finds that Spanish authors tend to be less clear, and that English authors use a greater range of lexis, such as ‘the preceding subsection’ in English as opposed to the more direct ‘*el apartado 2.5*’.

As discussed in Section 2.9, stance appears to be more frequent in English business management articles than Spanish (Orta *et al.* 2004), and Lorés-Sanz’s (2006) study of research article abstracts reveals that English demonstrates stronger position of the author through personal pronouns than Spanish. Similarly, there are more self-mentions in English than Spanish (Mur-Dueñas 2007b) and Orta (2010) studies stance taking in English and Spanish and finds that Spanish writers demonstrate stance very differently to English. In comparing writing in history in English and Spanish, many similarities have been found, with both languages using similar lexical bundles (Cortes 2008). Broader descriptions of Spanish academic writing argue that Spanish uses less participant-oriented metadiscourse in introductions, more subordination, more modals to create impersonality, and more passives than English (Mur-Dueñas 2007a). Reflecting on the established differences between English and French above, the reader may wonder how Spanish and French correspond in these respects. However, a limitation of these studies is the fact that it cannot be discerned whether languages that differ from English in similar ways are in fact similar to one another without also comparing those languages.

Returning to studies on Spanish academic writing, Lorés-Sanz (2011a; 2011b) finds that exclusive pronouns are more common in English than Spanish, but that they do correspond in terms of location where they are most used in the same sections in business management articles in English and Spanish in decreasing order of methods, introduction, results, and discussion. Other studies find that Spanish uses less hedging than English (Salager-Meyer 2011a) and that titles are longer in Spanish than English (Soler 2009; 2011). Again, here there is a clear binary presented in the literature that English is interpersonal and Spanish is not. Does this therefore indicate that French and Spanish are equally interpersonal? Not exactly, as the approach to contrastive analysis where only two languages are studied does not allow for a relative perspective across languages. Instead, studies of two languages only show one language in terms of another, often in a somewhat dialectical binary.

There are studies that have compared more than two languages such as Vassileva (2000) who studies authorial presence in English, German, French, Russian, and Bulgarian. This study shows that there appears to be a low degree of personalisation in French academic discourse (Vassileva 2000). In fact, it is not only low, but lower in French than English, German, Russian, and Bulgarian. Analyses of the research article conclusion sections in English, Spanish, and German indicates that more recent publications in Spanish demonstrate more dialogic conclusions than German, albeit to a much lesser extent than English (Ciapuscio and Otañi 2002). This is attributed to recognition within the literature that this change towards the interpersonal is a product of the globalisation of academia and the Anglicisation of academic writing. That is to say, contemporary Spanish has moved towards a more anglicised political correctness (Ciapuscio and Otañi 2002). Moreover, German does not reflect the same trend. As mentioned in Section 2.9, Salager-Meyer *et al.* (2003) and Salager-Meyer and Ariza (2004) find that French and Spanish have, traditionally, been more critical in academic writing than English and in book reviews Spanish is more critical than English and French. In their diachronic work on writer stance through criticism in English, French, and Spanish academic writing, they have also found that French and Spanish tend to be more direct, passionate, and authoritative than English which is typically more politically correct. Their study is a good example of comparative research that show French and Spanish as not only different to English, but to one another. They too, like Ciapuscio and Otañi (2002), find that

contemporary Spanish has moved towards a more anglicised political correctness where French seems, to a much lesser and slower degree, to be becoming more indirect. Elsewhere, the study of gender in determiner-noun relationships in Spanish, French, and English reveals important implications for English speakers learning French and Spanish (Hawkins and Franceschina 2004). Furthermore, Fløttum *et al.* (2006) have worked extensively on academic voices in English, French, and Norwegian in their KIAP corpus. Their work, contrasting a global language with a major and minor world language sees important differences in the use of voice in the research article in economics, linguistics, and medicine. In their research, English is seen to be more writer-responsible than French or Norwegian, as the use of metatext for guiding readers through writing is much more apparent in English.

A final project worth mentioning here is the EIIDA project¹³ which analysed academic discourse ‘to examine the impact of mode on the way scientific discourse is formulated and structured across languages (English, French, and Spanish) and disciplines (hard sciences vs. humanities)’ (Carter-Thomas and Jacques 2017, p.2). This project produced research such as Hartwell and Tutin’s (2017, p.57) work which finds that English uses more opinion verbs in spoken academic discourse and Hamilton and Carter-Thomas (2017) who find that French is more lexically diverse than English. Beyond these contrastive studies of English, French, Spanish and other languages there are more – see Suomela-Salmi and Dervin (2011) and Aijmer and Altenberg (2013). However, what has been discussed here is sufficient to exemplify the current state of corpus-based contrastive analysis. In the final part of this section, the focus turned to studies of three or more languages. These studies showed an increased depth and richness in terms of the claims they could afford as they did not see languages as binaries that are either similar or different to one another. This has been an important consideration for this thesis as will be seen in the following section, which addresses the contributions this thesis makes to the field of corpus-based contrastive linguistics.

¹³ Etude Interdisciplinaire et Interlinguistique du Discours Académique; Estudio Interlingüístico y Interdisciplinario del Discurso Académico; Interlinguistic and Interdisciplinary Study of Academic Discourse

3.6.2 Gaps and future directions in corpus-based contrastive linguistics

There are two important conclusions to draw from this brief discussion of the literature surrounding corpus-based contrastive analysis. The first pertains to the importance of relativism. Among the many valuable contributions evident in the review of contrastive research in Section 3.6.1, an evident dearth occurs as there are notably very few studies in contrastive linguistics on three or more languages (Van der Auwera 2012, p.75). In fact, in their edited volume on corpus-based contrastive linguistics Aijmer and Lewis (2017) define contrastive linguistics as the study of two languages side by side which reflects the general normalisation of contrastive analysis being almost exclusively constrained to two languages. However, as exhibited in Section 3.6.1, by comparing only two languages, the literature often presents findings in a binary fashion. That is to say, from a review of the literature comparing English and French, English appears to be interpersonal and French does not. Similarly, from a review of the literature comparing English and Spanish, English appears to be interpersonal and Spanish does not. Therefore, this binary view would position French and Spanish as similar when compared to English. This however, is a false equivalence. Research that exists that has focussed on three or more languages has made clear that languages when compared to one another are better positioned on a spectrum of relativity as opposed to a binary view. In such a spectrum, more languages allow for a greater understanding of the similarities and differences between them; for example, Spanish is increasingly reflecting English language norms in academic writing, and much more so than French.

That being said, there remain few studies that have compared more than two languages in a contrastive analysis. Interestingly, Van der Auwera openly ‘pleads’ (2012, p.69) for contrastive linguists to study three as opposed to two languages as he claims that including a third language can allow us to ‘make a point about the other two’ (Van der Auwera 2012, p.85). This thesis ascribes to this view and studies three languages in order to address this call. This PhD recognises that in studies with more than two languages, a clearer view of the relativity of interpersonality in each language can be seen. This is important in Section 7.2.9, for example, where in terms of question use, English appears more reader-oriented than French and Spanish while French appears more reader-oriented than Spanish. This contrasts heavily to the discussion of French, Spanish and reader and writer-responsible languages in Section 2.2.1. Overall, relatively speaking, these

languages are each quite different from one another in terms of interpersonality and the comparison of three languages herein demonstrates the importance of relativism and how languages' comparative behaviours are only understood in terms of one another.

The second important conclusion to draw from this review of research on corpus-based contrastive analysis discussed in Section 3.6.1 is the lack of attention generally paid to contrastive rhetoric theories of comparability. For example, the *tertium comparationis* rarely features in methodological descriptions of these analyses. Researchers like Connor and Moreno (2005) and Carter-Thomas and Rowley-Jolivet (2013) and those in ACTRES mentioned earlier in Section 3.6 are among the few contrastive studies found that directly refer to the *tertium comparationis*. Egan (2014), while not studying academic language, makes use of contrastive linguistic theories that allow him to account for directionality and relativism in his study of movement in French and English. However, overall, the two fields are not widely merged. Therefore, as a field there is still work to be done to draw together the theoretical foundations of contrastive and corpus linguistics. This thesis recognises this need and moves to merge the two fields more closely.

3.7 Conclusion

This chapter discussed the field of corpus-based contrastive linguistics and the theoretical and methodological foundations of that have informed this research. Endeavouring first to describe the field of contrastive analysis, this chapter addressed the negative perceptions of contrastive linguistics, discussed its evolution overtime, and identified the current state of the art in Section 3.2. Following that section, the core theories of contrastive linguistics were discussed in Section 3.3 and their application to contrastive linguistic methodologies was discussed in Section 3.4. These theories are integral to informing the methodology discussed in Chapter 4. Recognising this study as a corpus-based contrastive analysis, Section 3.5 discussed corpus linguistics, its evolution and theoretical underpinning, and its relevance to contrastive linguistics. Following that, the core studies in corpus-based contrastive analysis were discussed in Section 3.6 which make clear the current state of the field of corpus-based contrastive linguistics to which this thesis contributes as well as the gaps it addresses therein. The gaps identified refer largely to the fact that most studies in contrastive linguistics focus almost exclusively on two languages and few studies in corpus-based contrastive linguistics actively draw from

theoretical foundations in contrastive linguistics. This study, building on that research, draws from the theoretical concepts and methodological frameworks presented, as well as the current understanding of academic writing, reader engagement, and questions in English, French, and Spanish, discussed in Chapter 2, to conduct a contrastive study of academic writing in three languages. The following chapter, Chapter 4, presents the methodology employed to do so.

Chapter Four

A corpus-based contrastive analysis
of direct and indirect questions in
English, French, and Spanish

4 A corpus-based contrastive analysis of direct and indirect questions in English, French, and Spanish

4.1 Introduction

This chapter builds on the theoretical discussion presented in Chapter 2 and Chapter 3 and outlines the data and methodology applied in this study. First, in Section 4.2, the data are discussed. The English and French economics KIAP subcorpora are presented, followed by the comparable Spanish economics corpus which was created for this study. The process for the identification of questions within the data is also presented, indicating how direct and indirect questions were identified in the data. Next in Section 4.3, the *tertium comparationis* and equivalences that guide this study are presented, illustrating how questions are analysed in this study. Both *tertium comparationis* employed in this study are presented and with regard to equivalence, each one identified is outlined and linked to Krzeszowski's (1990) seven equivalences. Following that, in Section 4.4 and Section 4.5, the processes for conducting the contrastive analyses for which the results are presented in Chapter 5 and Chapter 6 are outlined. Finally, before concluding in Section 4.7, Section 4.6 discusses some of the data-related and methodological limitations of this study.

4.2 Data: KIAP-EEFS

Two thirds of the data in this research are taken from KIAP (Fløttum *et al.* 2006), a multilingual comparable corpus composed of 450 research articles, with 150 in English, French, and Norwegian respectively. These research articles are sub-categorised according to discipline with 50 in linguistics, economics, and medicine in each language. Permission was received from Prof. Kjersti Fløttum of the University of Bergen for the use of KIAP and this study used a subcorpus of research articles from economics in English (henceforth engecon) and French (henceforth frecon) and created a comparable Spanish subcorpus (henceforth specon) to conduct a trilingual study on KIAP-Economics, English, French, and Spanish – henceforth, KIAP-EEFS. Table 4-1 presents the size of each subcorpus and the total size of KIAP-EEFS when its subcorpora were analysed using

AntConc 3.5.8w (Windows)¹⁴ (henceforth AntConc). Note that AntConc was chosen as it is freely available online¹⁵ and allows for greater replicability by fellow linguists and teachers.

Table 4-1 Tokens in KIAP-EEFS

Subcorpus	Tokens in KIAP-EEFS
engecon	397,896
frecon	399,272
specon	412,631
Total	1,209,799

KIAP-EEFS is a small, specialised comparable corpus that contains texts ‘collected using the same sampling frame’ (McEnery and Xiao 2008, p.20). As such, it was important that the texts were equal in proportion according to language and genre and were collected within ‘the same sampling period’ (McEnery and Xiao 2008, p.20) to ensure the adequacy of the *tertium comparationis*. Taking a more in-depth view at ensuring this comparability, the following Sections 4.2.1 and 4.2.2 consider the composition and annotation of the three subcorpora.

4.2.1 KIAP: comparable subcorpus of English and French economics research articles

Fløttum *et al.* detail the construction of the KIAP corpus in their volume on academic voices (2006, pp.6-17). Here, the relevant details for the construction of engecon and frecon by Fløttum *et al.* (2006) are discussed, specifically for their role in guiding the construction of specon.

In the construction of engecon and frecon, a selection of journals that publish English and French language economics research articles were chosen based on the following criteria:

- These journals produced economics research articles that ‘contain reasonably large amounts of ordinary text in addition to the mathematical formulae that characterise most texts from this discipline’ (Fløttum *et al.* 2006, p.8).
- The journals represent quality publications and contain peer-reviewed and refereed articles.

¹⁴ Note that when searching engecon, the character encoding ‘unicode (UTF-8)’ is used while when searching frecon and specon, the character encoding ‘Western "Latin1" (iso-8859-1)’ is used

¹⁵ <https://www.laurenceanthony.net/software/antconc/> [last accessed 01 May 2019]

- These journals, where possible, are recommended by experts.

Upon collection, the engecon subcorpus contains 50 articles from the following journals/years:

- 18 articles from *European Economic Review* (2001–2003)
- 5 articles from *Journal of Economic Behavior & Organization* (2000–2002)
- 27 articles from *Journal of International Economics* (2000–2003)

Within which there are:

- 18 male authored papers
- 6 female authored papers
- 26 co-authored papers

Similarly, the frecon subcorpus contains 50 articles from the following journals/years:

- 34 articles from *Annales d'Economie et de Statistique* (1998–2001)
- 7 articles from *Economie Appliquée* (2000–2001)
- 9 articles from *Revue Economique* (2000–2001)

Within which there are:

- 25 male authored papers
- 7 female authored papers
- 18 co-authored papers

The journal articles in both engecon and frecon were only chosen if:

- they were written by native speakers representing their disciplinary discourse communities and cultural communities within the perspective languages;
- they were empirical or theoretical research articles *i.e.* not literature reviews or book reviews

In KIAP, each article is annotated and made available in AntConc readable XML files. The annotating process, as Fløttum *et al* identify, uses three types of tags to encode information (2006, p.12). Namely: extratextual; structural; and textual tags. Extratextual tags are used to give information about the text in terms of the author(s) and journal in which it was published. Structural tags are used to identify the individual sections of the article such as abstracts or introductions *e.g.* <abstract> text </abstract>, which identifies the abstract as the text between the two tags. Swales' IMRD (1990; 2004) is less apparent in the articles in KIAP and in KIAP-EEFS all articles are coded as <intro> <mid> and <concl>, where <mid> refers to the 'thematic middle sections' (Fløttum *et al.* 2006, p.12). Finally, textual tags (Fløttum *et al.* 2006, p.14) are used to identify subtitles, quotes, tables, and examples. Each type of tag outlined here is equally important as they are used

in identifying where in the research article the searched items occur and which elements of the text are to be included for analysis. An important piece of metadata worth noting is that only 48 texts in English and 43 texts in French were found that have explicit conclusion sections. Where no explicit conclusion sections are present, the article proper is closed with the </mid> tag.

4.2.2 Building a corpus of Spanish economics research articles comparable to KIAP

The previous discussion of both *engecon* and *frecon* and their collection, composition, and annotation by Fløttum *et al.* (2006) has been replicated in the collecting and building of *specon*.

Following Fløttum *et al.* (2006), Spanish/Hispanic experts in economics were contacted to identify journals they considered to be quality publications based on a short-list of peer-reviewed journals. The following journals were selected and permission for their use requested¹⁶:

- *Investigación Económica* (Mexico)
- *La Revista de Economía Mundial* (Spain)
- *El Trimestre Económico* (Mexico)

Upon receiving permission, a database of every article published in each journal between the years 1999 and 2003 that was available in a digital format was created. Following Fløttum *et al.* (2006), articles were excluded based on the following criteria:

- all texts that were not empirical or theoretical research articles were not chosen *i.e.* literature reviews or book reviews;
- articles where none of the authors are native Spanish speakers were not chosen¹⁷.

After the delimitation of articles for selection was complete, a database of articles was created for each journal according to the year and authorship (male, female, and co-authored papers). Following Fløttum *et al.*, efforts were made where possible to have a distribution of authorship across journals as they state that ‘we have attempted to ascertain

¹⁶ The permissions were for the use of research articles in the development of research data for use in this PhD and further research. In the case of each journal, the journal maintained copyright of each publication. Therefore, individual authors need not have been contacted for permissions.

¹⁷ Only articles where first language speakers were chosen in order aim for as culturally homogenous group as possible. The goal is to ensure that the texts are not culturally skewed by writers from different academic traditions and backgrounds.

that both single-author and multi-author articles are represented, and, for the single-author ones, articles written by women as well as by men' (2006, p.9). In the case of specon, it was decided that of the 50 research articles, they would be distributed as evenly as possible with 17 from both *Investigación Económica* and *La Revista de Economía Mundial* and 16 from *El Trimestre Económico* as this was possible within the data for which I received permission to use. The goal was to choose an even distribution from each year, and within that, from male and female authored papers and co-authored papers. Where this was not possible at times, it was decided that an equal balance of male and female authored papers and co-authored papers was more important than an equal number from each year as the articles are written over quite a short period of time and likely cannot account for diachronic variations. It was also decided to maintain an equal male/female distribution and have more co-authored papers to balance the gender representativeness. Finally, the eligible articles were each attributed numbers, and using a random number generator function in Microsoft Excel were chosen and included in the specon subcorpus. A full list of articles included in specon is available in Appendix A. Table 4-2 outlines the distribution of the chosen articles within each journal. Note that "M" signals male authors, "F" signals female authors, and "C" signals co-authored papers.

Table 4-2 specon composition

Journal	1999			2000			2001			2002			2003			Total			
	M	F	C	M	F	C	M	F	C	M	F	C	M	F	C	M	F	C	All
Investigación Económica	1	1	2	1	1	2	1	2	1	1	1	1	1	0	1	5	5	7	17
La Revista de Economía Mundial	0	0	0	1	1	2	1	1	1	1	2	1	2	1	3	5	5	7	17
El Trimestre Económico	1	2	1	1	0	2	1	1	1	1	1	1	1	1	1	5	5	6	16
Total	2	3	3	3	2	6	3	4	3	3	4	3	4	2	5	15	15	20	50

The articles were converted to Microsoft Word and checked for accuracy, accounting for any errors that may have occurred owing to file conversion. Subsequently, following the collection and construction of engecon and frecon, specon was annotated and converted

to AntConc readable XML files. Regarding this annotation, it was found that not many articles in Spanish contain conclusion sections. In fact, on 39 of the 50 texts in specon contain explicit conclusion sections. This directly contrasts to English and French. The subcorpus specon was manually tagged following the tagging and annotation process discussed in the Section 4.2.2. Following the construction of specon, KIAP-EEFS was ready to be searched for questions.

4.2.3 Finding questions in KIAP-EEFS

The primary focus of this research is questions that form part of the authorial text and that serve to render the research article as interpersonal *i.e.* questions allow the author to communicate and interact with the reader. Only those items that occur in between the opening tag <title> and the closing tag </title> and the opening tag <intro> and the closing tag </concl>¹⁸. This study excludes questions within examples, tables, and quotations as they were not considered interactive in terms of writer/reader interaction; that is to say, only those questions, included by the author to pose a question in order to interact with the reader in some way were included. Based on the discussion of questions in Section 2.10.2, this study included the analysis of both direct and indirect questions.

Direct questions are quite easily searchable, and can be identified by searching for the question mark character as the question mark is an illocutionary force indicating device (Flöck and Geluykens 2015, p.7) which acts as an accessible ‘starting point for electronic searches in language corpora’ (Flöck and Geluykens 2015, p.7). However, identifying indirect questions was much more challenging. What follows is a documentation of the identification process for indirect questions.

1. In each language all instances of wh-complements, relative pronouns, if conjunctions, and questions words were searched *e.g.* what, *qu’est-ce* and *qué*.
2. Given the quantity of items, sample concordance lines of each KWIC were captured. To generate appropriate samples, a sample size calculator¹⁹ was used and each sample was extracted at 95% confidence +/- 5%. This means that there is 95% confidence that the sample is balanced.

¹⁸ Where there is no conclusion section, the search was confined to the closing tag </mid>

¹⁹ The sample size calculator used is available: <https://www.surveysystem.com/sscalc.htm> [last accessed 15 June 2019]

3. Each sample was analysed to identify indirect questions.
4. As indirect questions can occur without such wh-complement, the search could not be constrained to this form. Therefore, the words used in conjunction with these questions words to raise indirect questions (identified in step 3) were noted *e.g.* ask, *se demander*, *cuestionar*.
5. To identify indirect questions and not constrain by form:
 - a. All verb, adverb, adjective, and noun forms related to the identified words were searched and captured *e.g.* question as a noun (question), verb (question), adjective (questioned, questionable), and adverb (questionably) in each language.
 - b. Other words known to be used to raise indirect questions drawn from previous research were searched and captured *e.g.* question words in Table 2-1. Furthermore, through a close read of the frequency lists, and words presumed to have the potential to indicate a question were searched.
 - c. Confidence samples were extracted once again from all KWIC searches at 95% confidence +/- 5%.
6. These samples were analysed and all indirect questions were extracted.
7. Finally, any questions, direct and indirect, not deemed to be part of the authorial text were excluded.

This resulted in the identification of the following number of questions in KIAP-EEFS as presented in Table 4-3. Note that here there is no distinction between direct and indirect questions at this time. As question type is a variable, it is addressed in Chapter 5, Section 5.2.2.3.

Table 4-3 Question in KIAP-EEFS

Language	English	French	Spanish
Total	201	129	98
WPM	505	323	237

This identification of indirect questions was important as the only assumption permissible in a contrastive analysis is the *tertium comparationis* which is tested by equivalences such as form. As such, if only wh-complements were studied, then all indirect questions would likely be content questions which would reflect the methodological circularity discussed in Section 3.4. Instead, adopting the above-outlined methodology, this research has taken

a very broad approach to capturing indirect questions which will allow the questions to be analysed in terms of form.

4.3 *Tertia comparationis* and equivalences: comparing and contrasting questions found in KIAP-EEFS

Drawing on the theory discussed in Chapter 2 and Chapter 3, this section describes the *tertia comparationis* and equivalences used in this corpus-based contrastive analysis.

4.3.1 *Tertia comparationis*

In Section 3.3, it was discussed that contrastive analyses involve identifying and testing an assumed similarity between languages. This assumed similarity is a *tertium comparationis* which forms the basis of a contrastive analysis. Typically, a *tertium comparationis* is retested using different criteria in order to evaluate the assumption from different angles. In the case of this study, the *tertium comparationis* is retested eight times by 8 equivalences. In theory, this pattern could continue with further criteria. In this study, two analyses are presented requiring the following two *tertia comparationis*:

1. The first *tertium comparationis* identified is the assumption that English, French, and Spanish economics research articles contain questions used as a rhetorical feature of reader engagement. The results of the analysis of this *tertium comparationis* are presented in Chapter 5.
2. The second *tertium comparationis* identified after the first analysis is the assumption that English, French, and Spanish economics research articles share a rhetorical functional use of questions as reader engagement. The results of the analysis of this *tertium comparationis* are presented in Chapter 6.

In theory, the findings of Chapter 6 could issue a further test and analysis of a *tertium comparationis*. However, it is beyond the scope of this project to continue down this line of reasoning in this thesis.

4.3.2 Equivalence

The equivalence used to test the *tertium comparationis* in Chapter 5 and Chapter 6 were chosen as they emerged in the analysis of questions, as discussed in Section 1.3.3. There are many further possible equivalences that could be used to test these *tertia comparationis*. However, this study has constrained the analysis to the study of equivalences in the form of question frequency, question function, question length, question type and form, question location in the research article, passivity in questions,

tense of questions, and sentence type within questions. In each case, the criteria for analysis were defined broadly to avoid anglocentrism. How these were analysed will now be briefly presented.

4.3.2.1 Statistical equivalence: question frequency

Statistical equivalence reflects question frequency which is determined through AntConc which counts the occurrence of searched queries. These frequencies are presented in words per million in this study in order to normalise for fair comparison across languages. They are words per million and not questions per million as the frequency is relative to the size of the corpus data, which is counted in words.

4.3.2.2 Pragmatic equivalence: question function

Pragmatic equivalence here is linked to the function of questions, which, in this study, borrow from and build on Hyland (2002a). The explanations of these functions are drawn from the description in Chapter 2, Section 2.10.1 and Table 4-4 presents a reminder of each function and their remit.

Table 4-4 Seven functions of questions as reader engagement

Question function	Function outline
Getting Attention and Focussing the Reader	These questions refer only to those used in titles of research articles to attract a reader's attention
Framing the Discourse	Questions are used to inform the reader of the questions that are guiding the research; these are questions that the research intends to (try to) answer
Organising the Text	Questions are used as in-text signposts to guide the reader; questions are used in subheadings as a structuring strategy
Creating a Niche	Questions are used to create a niche and identify a gap
Expressing an Attitude and Counter-claiming	Questions are used to critically evaluate works by other researchers with which a writer does not necessarily agree
Setting up Claims and Protecting the Writer	Questions are used to express an opinion or create an element of doubt which allows the writer to express stance or attenuate doubt with a claim
Asking Real Questions and Pointing Forward	Open questions that have no anticipated response; questions that maintain a rhetorical function of self-promotion as an identifier of problems; questions that look beyond the text to future problems and considerations

Overall, Hyland's functions were modified in order to incorporate further research from the studies of academic discourse and questions, discussed in Chapter 2, Section 2.10.1. As discussed in Section 2.10.1, the decision to use the modified version of Hyland's framework emerged from its successful application in a pilot study, wherein the framework captured all the question functions identified in the data. One important note in the application of these questions to languages other than English was the need to ensure that the function is not constrained by other criteria like location. As such, no function was seen to predict or be predicted by location except for the function: getting attention and focussing the reader which is defined as questions used in titles. In the case of this function, there is no location related finding, in order to avoid methodological circularity. These functions have been successfully applied to this study and no other functions of questions were identified in the analysis.

Attributing a function to the question is a qualitative endeavour and as such it is often necessary to look at context as well as the cotext to understand the functions. In this study, this has involved looking within the research articles to see the role the questions play therein. While this is quite time-intensive, the smaller nature of the data set means that it is possible to take a more in-depth analysis of the data which might be challenging in larger corpora.

4.3.2.3 Rule equivalence: question length

Rule equivalence is concerned with sentence type and length. Question length is determined by counting every word in the sentence, not just lexical words. Question lengths are calculated in terms of average, median, and mode²⁰ length as well as in a range of words from 1-15, 16-25, and 26 words or more to present both nominal and categorical data.

4.3.2.4 System equivalence: question type and form

System equivalence governs form-related equivalences and relates here to question type and form. Questions in this study are one of two types: direct or indirect. Direct questions are identified by a question mark. Indirect questions, following the discussion in Section

²⁰ Mode in this case refers to the most frequently occurring question length. For example, if there are 20 sentences of 10 words exactly and there are no other lengths that occur more than 20 times, then the mode would be 10.

2.10.2, are embedded questions which may be raised by either a *wh*-complement or are concealed questions. Concealed questions are those which do not contain an interrogative device but that serve to signal information that will be revealed. As presented in Section 2.10.2, indirect questions do not need to have a *wh*- complement, they simply need to have the intention to reveal some information (Romero 2005).

As indirect questions can be categorised according to the same forms as direct questions (Suñer 1993), all questions were then categorised as either content, polar, declarative, alternative, and elliptical questions. Content questions require information as an answer, polar questions require a yes/no answer, declarative questions are syntactically accurate affirmative statements, alternative questions give a choice of answer, and elliptical questions elide some words from the formation of the question. It is of note that although Section 2.10.2 problematises the designation of elliptical questions as a question form, it was deemed important here to capture ellipsis given that such a syntactic feature is typically expected in spoken and less formal registers. For direct questions this is clear from the formal presentation; however, for indirect questions the form may be concealed and linked to the answer where if the answer is information, it is an indirect content question and if the answer is yes or no, it is an indirect polar question. Overall, these labels were chosen to avoid anglocentricity in question form.

4.3.2.5 Translation equivalence: question location in the research article

Location in the research article falls under translation equivalence as it refers to the generic construction of research articles across languages and is a cultural phenomenon within the discourse communities. It is governed by the structural tags used: <title>, <intro>, <mid>, and <concl> which refer to text within the titles, introduction sections, middle sections, and conclusion sections, respectively.

4.3.2.6 System equivalence: passivity in questions

System equivalence governs form related equivalences and relates to passivity. To identify passivity, the passive voice is used in English, French, and Spanish. Identifying passivity in French and Spanish also allowed for the reflexive passive and *on* passive in French and the *se*-passive in Spanish, as discussed in Sections 2.6.3 and 2.7.2. Passivity or lack thereof was determined from the analysis of the KWIC word or question verb used to raise the direct and indirect questions studied.

4.3.2.7 System equivalence: tense and use of verbal modality in questions

System equivalence governs form-related equivalences and relates here to tense and verbal modality. Tense is categorised broadly as past, present, future, no tense, or verbal modality. While English contains no future tense, but a future aspect, and contains many past and present aspects, French and Spanish both contain future tenses and several past tense. For simplicity, tense is used herein for all three languages, although it should be noted that tense includes both tense and aspect in each language. Overall, there is no specific break down of individual past and present tenses as it was too granular for the data studied herein. For verbal modality both modal verbs and the subjunctive were used to assign modality to questions. Owing to the complexity of question sentences in the data, tense or modality thereof was determined exclusively from the analysis of the KWIC word or question verb used to form the direct and indirect questions studied. It should be noted that as tense predicts a verb form as question raising, it is limited where IFIDs were not verbs.

4.3.2.8 Rule equivalence: sentence type within questions

Rule equivalence is concerned with sentence length and type. For sentence type, this study takes a general and easily comparable view, distinguishing between simple, complex, compound, and compound-complex sentences. Such a view takes simple sentences as those with one clause, complex sentences as those with subordinating clauses, compound sentences as those with coordinating clauses, and compound-complex as those with both coordinating and subordinating clauses (Carter and McCarthy 2006). In fact this description of sentence composition has long been established in English (Huddleston 1984), reflects Spanish language descriptions (Revilla 1980; Alarcos Llorach 1999), and though often labelled differently, captures French use of coordination, and combinations of coordination and subordination in the composition of sentences (Collins 2011; Hawkins and Towell 2015).

4.3.3 Summary

Overall, the two *tertia comparationis* analysed in this study are analysed in terms of the eight equivalences presented in this section. This of course excludes the equivalence for which they account *i.e.* pragmatic equivalence and function in Chapter 6 account for the *tertium comparationis*, so it cannot be tested in terms of function. The following section

outlines the step by step process of the corpus-based contrastive analysis methodology wherein the questions identified in KIAP-EEFS are contrasted in English, French, and Spanish based on the *tertium comparationis* and identified equivalences.

4.4 Corpus-based contrastive analysis: question use as a shared rhetorical feature of reader engagement in English, French, and Spanish

The first corpus-based contrastive analysis presented in this thesis is a convergent contrastive analysis which centres on the following *tertium comparationis*: the English, French, and Spanish economics research articles contain questions used as a rhetorical feature of reader engagement.

As it is a convergent analysis, questions are identified independently in each language and then compared with one another. These questions are identified in KIAP-EEFS and are tested in terms of the equivalences discussed in Section 4.3.2. Figure 4-1 illustrates this process.

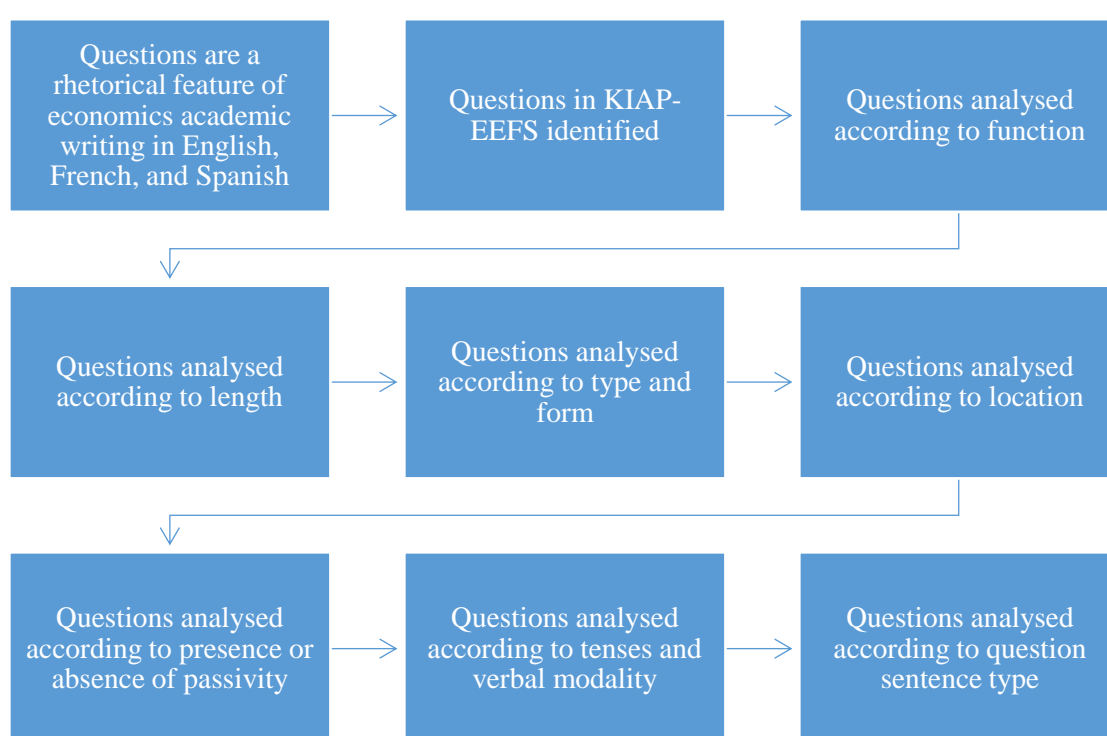


Figure 4-1 Contrastive analysis of questions as a shared rhetorical feature of reader engagement in English, French, and Spanish

The results of this analysis presented in Chapter 5 present a number of useful insights. First they give a general view of the use of questions in economics academic writing in

English, French, and Spanish. They describe the relative frequencies of different features of questions within and across languages. They identify how questions as a rhetorical feature of reader engagement correspond in English, French, and Spanish. Finally, they offer a benchmark from which to compare the results of Chapter 6 which takes shared function as *tertium comparationis*. The process for this second analysis is outlined in the following section.

4.5 Corpus-based contrastive analysis: question functions as a shared rhetorical feature of reader engagement in English, French, and Spanish

The second corpus-based contrastive analysis presented in this thesis is also a convergent analysis which centres on the following *tertium comparationis*: the English, French, and Spanish economics research articles share a rhetorical functional use of questions used as reader engagement. Again, as it is a convergent analysis, questions are categorised independently in each language and then compared with one another. These questions are categorised according to the seven functions of questions identified in Section 4.3.2.2. Questions in each function are tested in terms of the equivalences discussed in Section 4.3. Figure 4-2 illustrates this process.

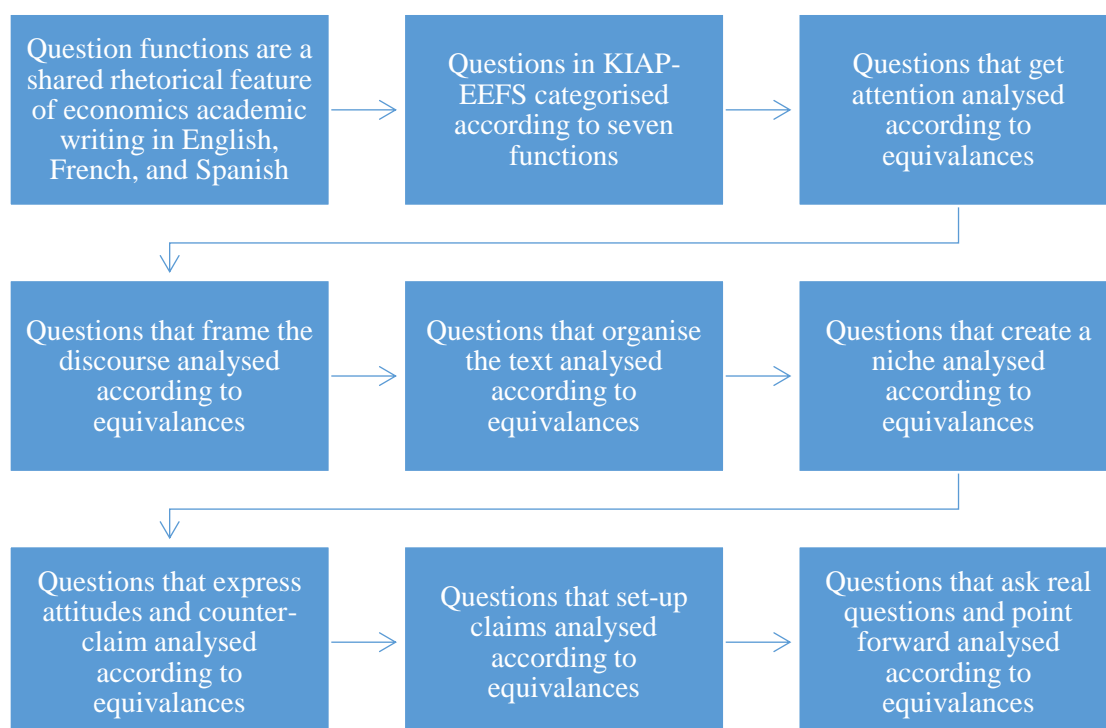


Figure 4-2 Contrastive analysis of questions with a shared rhetorical functional use as reader engagement in English, French, and Spanish

The results of this analysis presented in Chapter 6 offer a number of useful insights. First, they give an alternative view of question use to Chapter 5, where both frequent features of functions as well as salient features of functions in economics academic writing in English, French, and Spanish are identified. They identify how each function of questions as a rhetorical feature of reader engagement corresponds in English, French, and Spanish in terms of frequency and saliency, where frequency refers to what is most common in each function and saliency refers to what is unusual and distinctive about each function.

An important item to note here is that findings are referred to as unusually frequent or unusually infrequent throughout Chapter 6. This evaluation is determined from the comparison of the frequency of a feature within a question function to the overall presence of that feature in the same language data. This means that, for example, if a function like getting attention and focussing the reader contains 4.5% of questions, we might expect a 4.5% distribution of equivalences for this function if the function has no relationship to the variables. However, in English, this function contains 9.1% of questions that contain fewer than 15 words in English. Therefore, the dispersion is not normal and this function contains an unusually frequent use of short questions when compared to overall patterns of use. This is because an normal distribution would predict 4.5% of short questions occurring within 4.5% of questions; however, the unusually frequent use of short questions to get attention and focus the reader renders short questions a notable feature of the function. Such a perspective allows for the identification of not only frequent results, but salient results where proportionally, a function may contain more examples of a feature of questions than its proportion of questions would predict. This distinction is important for understanding the results presented in Chapter 6. The following section outlines some of the limitations of this study before the results are presented in the next chapter.

4.6 Identifying and addressing limitations

Limitations in this study have been defined as relative to data or methodology. These are now briefly discussed.

4.6.1 Data limitations

This section delineates some of the limitations surrounding the data used in this study while also offering an explanation as to how the impact of these limitations has been minimised.

4.6.1.1 KIAP-EEFS

It is important to note that, relative to general corpora such as the BNC or the Cambridge English Corpus, KIAP-EEFS is small in stature. Although the corpus amasses a total of 1,209,799 tokens, as monolingual subcorpora, each subcorpus constitutes a small sample which limits this research from making many claims of statistical significance beyond the samples studied. However, following Aston (2001), this is not problematic, as this is often the nature of small-specialised corpora which can overcome weaknesses in size through their representation of specialised and heavily patterned language. Moreover, the small size has allowed for the in-depth analysis of the data which would have required random sampling or reduced scope in a greater data set. However, owing to the smaller sized corpus, there are some question functions in Chapter 6 with very little data *e.g.* expressing an attitude and counter-claiming in Section 6.6. As such, it is difficult to draw wider-reaching conclusions for those functions.

Furthermore, the corpora studied here in each language are confined to economics and, owing to limitations of space, are not compared to broader reference corpora, corpora from different disciplines, or corpora composed of different genres of academic texts. While of course it would be interesting to compare such data sets, this is considered to be beyond the scope of this study and instead is considered a potentially valuable follow up study.

Finally, the impact of the publishing process and the specific style of journals has not been captured here. As these data were published between 1995 and 2003, it is not possible to access the guidance given to writers who published at this time in these journals. Therefore, it is not possible to see any relationship between the specific journal requirements and the writing produced in the research articles. This would be valuable metadata were it available, and in future studies, it would be worth capturing such data in the construction of an academic corpus.

4.6.1.2 Author identity in KIAP-EEFS

Fløttum *et al.* acknowledge that the identification of both the authors' genders, and that the authors are native speakers, and therefore representative of their discourse community, is problematic as these demographic details cannot be definitively known (2006, p.10):

An author's nationality and mother tongue cannot be decided on with certainty on the basis of his or her name. Thus, it is possible that there are a couple of corpus articles written (only) by non-native authors. We do not believe, however, that the number of these can be large enough to jeopardise representativity.

Similarly, they acknowledge that the inclusion of co-authored research articles in their data can be problematic, particularly where articles are co-written by authors from differing linguistic backgrounds. In the case of Fløttum *et al.*, only co-authored papers wherein one author could be identified as a native speaker of the language in question were chosen. However, 'articles with several authors of which one (or more) is a non-native speaker of the article language may well have that very person as its main author' (2006, p.10). Moreover, they recognise that it cannot be known what impact the process of article publication has had on the research articles in terms of editors and reviewers' input, and therefore, what is solely representative of male, female and co-authored papers. It should be noted that the reasons to focus on native speakers relates to their study of culture, where nativeness is used as a proxy for determining nationhood and culture. This study similarly uses culture as a means to interpret the findings of this study and therefore makes similar constraints.

Following Fløttum *et al.*, it is believed that these limitations are not so substantial that they would 'jeopardise representativity' (2006, p.11) or compromise the value of the results. For the purpose of their study, and this study, the research article as a communicative event is not necessarily representative of one author's writing practices. Rather, it is seen here as the accepted standard of a discourse community which includes the impact of editors and reviewers on the writing, raising it to an international standard of academic scientific writing, and not that of individual authors alone. Owing to this international, peer-reviewed standardisation of research articles, it is unsurprising that Fløttum *et al.*, in their work on academic writing, found 'haphazard' (2006, p.265) differences in female and male writing that, for them, did not represent any reliable tendencies.

4.6.1.3 Comparability challenges in KIAP-EEFS

Finally, issues surround the comparability of the data must be considered. Each subcorpus is composed of research articles from three journals and contains male, female, and co-authored articles. The distribution of articles between and within each journal in each language differs as identified in Section 4.4.2. Fløttum *et al.* (2006) identify that many of the research articles in the French subcorpus come from one journal due to limitations of accessibility. However, knowing that as these journals are seen to be representative of the discourse communities where the writing is peer-reviewed, refereed, of an international standard, and centres on general economics, these uneven distributions can be considered unproblematic. Moreover, as mentioned in the previous section, although there is an uneven representation of male, female, and co-authored papers in engecon, frecon, and specon, this is not problematic, as Fløttum *et al.*, in their work on academic writing, found ‘haphazard’ (2006, p.265) differences in female and male writing that, for them, did not represent any reliable tendencies. Furthermore, here the focus is on the writing of the disciplinary discourse communities as a whole and there is no explicit focus on gender.

4.6.2 Methodological limitations

This section outlines the challenges and limitations of the contrastive analysis approach taken in this research.

4.6.2.1 Limitations of contrastive analysis approaches

A limitation of the approach of this study is linked to contrastive analysis. In contrastive analyses, the claims that can be made are limited and there is a need for constant retesting. This is due to the fact that an effective contrastive analysis necessitates one *tertium comparationis* which is tested in terms of many equivalences. This means that it is not possible to see the relationships and correlation between different equivalences *i.e.* there is no way of knowing, from the results presented in Chapter 5 and Chapter 6, if there is any salient relationship between passivity and question form, for example. However, future tests could place passivity as the *tertium comparationis* which could then be tested.

Finally, in order not to mislead the contrastive analysis, equivalences used here required broad definitions. As such, there is no detailed consideration of the individual tenses, the functions of passives and modality, or further equivalences. This would be too granular

in data of this size. However, future studies can elaborate further on some of these equivalences.

4.6.2.2 Translation competence and judgement variables in corpus-based contrastive analysis

There has been a lot of controversy surrounding what constitutes a correspondence across languages as researchers have identified that their analyses are based, to a large degree, on bilingual (Krzeszowski 1990, p.161) or translation competences (Chesterman 1998, p.38) requiring the contrastive researcher to make an informed judgement (Chesterman 1998, p.16).

Contrastive linguists make these judgments based on their language competence and expertise and for Chesterman, the role of translator competence is unavoidable (Chesterman 1998, p.16). In many contrastive studies, such as Fløttum *et al.* (2006) or Salager-Meyer (2011a), the role of judgement is very much apparent. However, for Chesterman (1998), the retesting of results can account for this, as the results of such analyses can continue to be tested in different ways and by different means in order to validate findings. The interpretative nature of a contrastive study is linked to the functional focus where ‘illocutionary force indicating devices’ (Flöck and Geluykens 2015, p.7) for questions are identified in a function-to-form approach to corpus linguistics. This involves the analysis of question marks and adjectives, nouns, and verbs which act as illocutionary force indicating devices for indirect questions, identified through the process outlined in Section 4.2.3. These questions are then analysed in terms of their function, looking beyond the vertical study of language in abstraction in concordance lines, but to a combination of corpus vertical-reading and horizontal-reading which allows for a focus on language in communication, the cotext and the context (Aijmer and Rühlemann 2014, p.9). To support this approach and avoid subjectivity, the convergent analysis can help address some of these issues, as this involves identifying questions in each language and then comparing. This means that the comparison of like with like is ensured where likeness here is defined by the item studied being a question.

Typically to address judgement variables, inter-rater variation is checked in order to overcome misinterpretation. Unfortunately, this was not possible in this study as it involves three languages and multiple variables that require in-depth qualitative analysis

of questions and their texts. Instead, a form of intra-rater variation was checked instead. This means questions were analysed in each language and later reanalysed to see if any changes needed to be made. As such, variation has tentatively been accounted for, though in an ideal research context this would be a joint project to allow for inter-rater correspondence checking.

4.7 Closing remarks

This chapter has outlined the corpus-based contrastive analysis methodology employed in this research as well as the data to which it is applied. Following consideration of the data and the composition of KIAP-EEFS, the identification of relevant engagement markers within these data was outlined and discussed. The analytical framework used to contrast and analyse the questions in English, French, and Spanish was subsequently considered, delineating the shared *tertia comparationis* that form the basis of Chapter 5 and Chapter 6. Subsequently, the process of comparing and contrasting these engagement markers was presented, and finally, the limitations of this study were presented in terms data and methodological limitations. Moving forward, the next two chapters will present the results of this study. Chapter 5 presents the results of the corpus-based contrastive analysis of question use as a shared rhetorical feature of reader engagement in English, French, and Spanish, and Chapter 6 presents the results of the corpus-based contrastive analysis of question functions as a shared rhetorical feature of reader engagement in English, French, and Spanish.

Chapter Five

Corpus-based contrastive analysis of question use as a shared rhetorical feature of reader engagement in English, French, and Spanish

5 Corpus-based contrastive analysis of question use as a shared rhetorical feature of reader engagement in English, French, and Spanish

5.1 Introduction

This chapter presents the findings of the first of two corpus-based contrastive analyses of questions in economics research articles in English, French and Spanish. In this chapter the *tertium comparationis* identified is the assumption that English, French, and Spanish economics research articles contain questions used as a rhetorical feature of reader engagement. The goal is to test this assumption based on the equivalences detailed in Section 4.3.2. To this end, first, Section 5.2 establishes to the *tertium comparationis* of the contrastive analysis by demonstrating that questions are a comparable feature of academic writing across KIAP-EEFS. Following that, Sections 5.3-5.7 tests this *tertium comparationis* by comparing questions in each language in terms of function, length, type and form, location, and passivity, tense and verbal modality, and sentence type. Based on this, Section 5.8, presents the manifest similarities across the three languages and Section 5.9 proposes a correspondence for questions in English, French, and Spanish. This is followed by a brief conclusion in Section 5.10.

5.2 *Tertium comparationis*: question presence in KIAP-EEFS

As seen in Section 4.2, questions were identified in KIAP-EEFS using various search methodologies. Each subcorpus was found to contain questions across a range of texts and the exact figures are presented in Table 5-1²¹:

Table 5-1 Questions in KIAP-EEFS

Language	English	French	Spanish
Total	201	129	98
WPM	505	323	237

²¹ For efficacy of comparison, all figures presented after Table 5-1 are normalised to words per million and rounded to whole numbers. This study is contrastive in nature and therefore necessitates a fair comparison which normalisation affords. Moreover question relative frequencies are presented per million words as the frequency is based on the overall size of the subcorpora which are counted in words.

Questions within each subcorpus have a relatively normal dispersion. Figure 5-1 shows that the mean (red bar) and median (black bar) are relatively comparable and not overly skewed by outliers, which are the blue dots above the box plots. This is further bolstered when a Juilland's D deviation test was applied to question dispersion and the English, French, and Spanish subcorpora scored highly at 0.83, 0.84, and 0.81 respectively²².

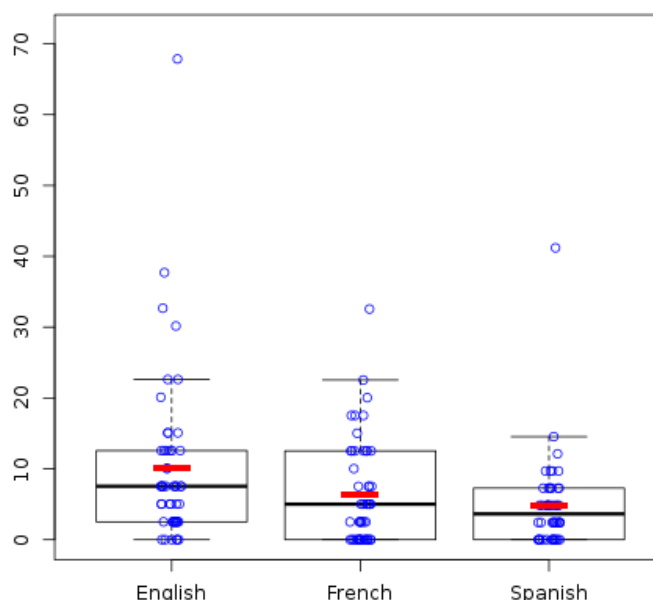


Figure 5-1 Dispersion of questions in KIAP-EEFS in WPM

Overall, questions are a feature of all three subcorpora and therefore worthy of study which is made clear in Table 5-1 and Figure 5-1. However, questions are used to varying degrees in each language with English using them more than twice as often as Spanish and with French situated in between the two. In fact, while there is no significant difference in question use between English and French and French and Spanish, there is a significant difference between English and Spanish, as reported in Figure 5-2.

²² A score of 1 Juilland's D deviation test means there is a perfectly even distribution. Therefore, a score closer to 1 than 0 has a more even distribution. These calculations and Figures 5-1 and 5-2 were generated using Lancaster Stats Tools Online (Brezina 2018)

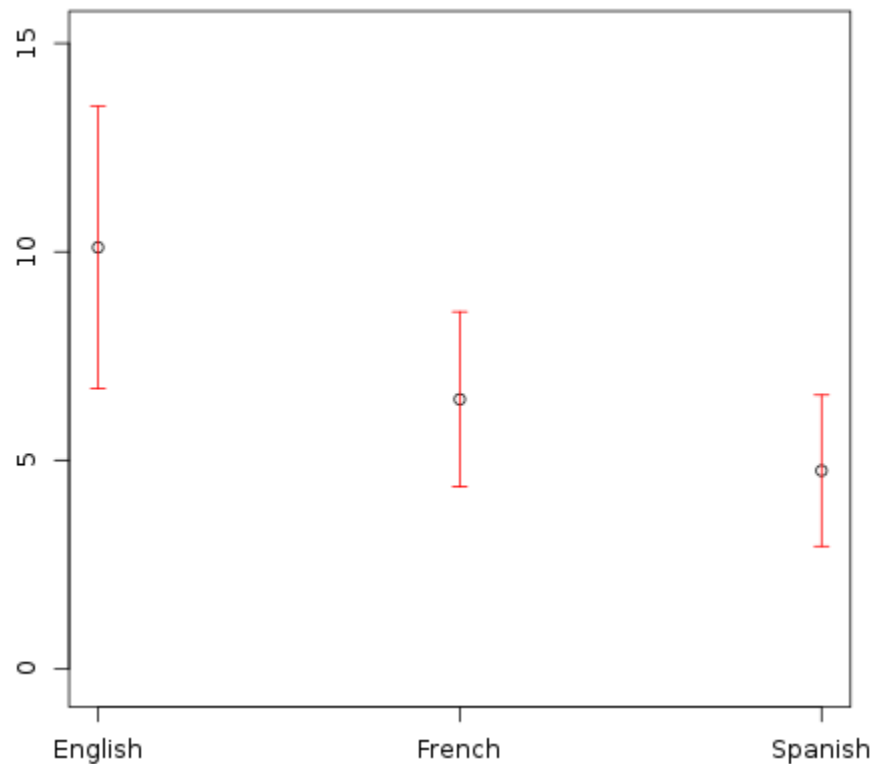


Figure 5-2 95% confidence intervals for question use across languages in WPM

Spanish uses significantly fewer questions than English which is evidenced in the confidence intervals which are the red bar plots in Figure 5-2. Overlapping bars, like English and French, and French and Spanish indicate that there is no significant difference in their use in the relevant data (Brezina 2018, pp.31-32). However, non-overlapping bars like English and Spanish indicate a significant difference in the presence of questions within the subcorpora (Brezina 2018, pp.31-32). Given that both the dispersion test in Figure 5-1 and the Juilland's D deviation scores mentioned earlier indicate a relatively normal and comparable dispersion, the significant difference is not the result of outliers. This evidence suggests that questions are a notable and shared rhetorical feature of economics academic writing in each language. Moreover, it suggests that Spanish uses significantly fewer questions than English. Therefore, questions emerge as a feature that is worthy of comparison.

Given that questions are a rhetorical feature of academic writing in each language, they are established as a *tertium comparationis*. However there is an indication already of differences across languages in terms of frequency. To further identify similarities and

differences in question uses, the following sections test this assumed similarity in terms of function, length, type and form, location, passivity, tense and verbal modality, and sentence type.

5.3 Question functions

The functions into which questions are categorised were established in Section 4.3.2.2. Each language tends to use questions for the same functions to a similar degree with some variations. The distribution of these functions is similar for each language as presented in Figure 5-3.

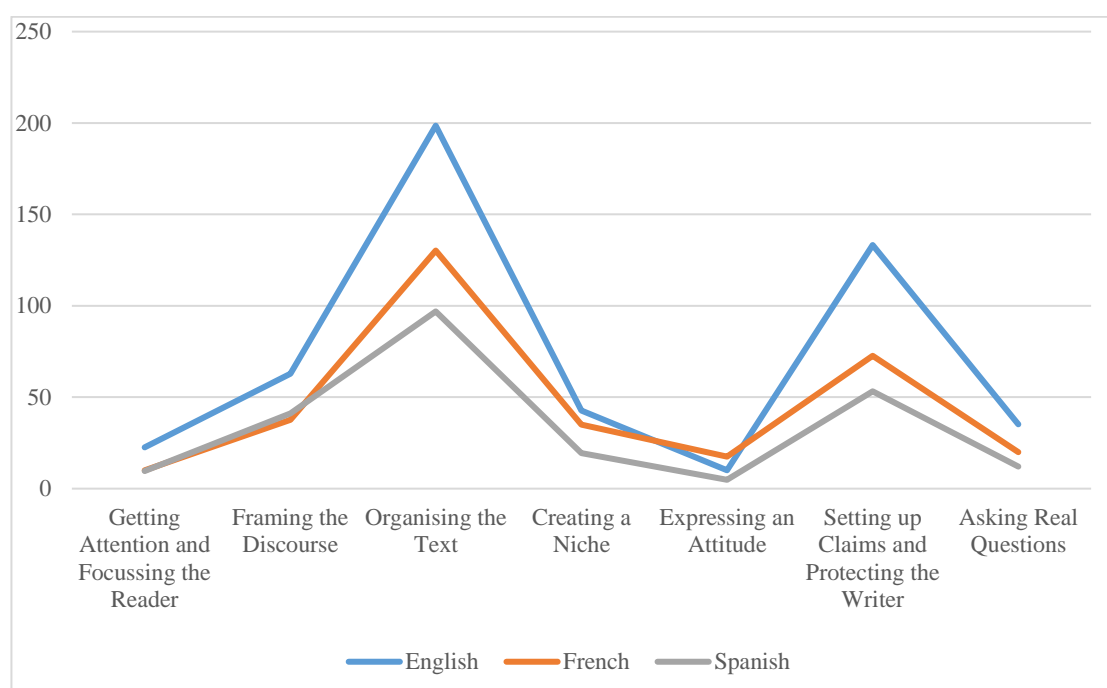


Figure 5-3 Distribution of questions per function in WPM

Within each language, there are proportional differences, as the relative frequency of some functions is higher in different languages. In order to give a clearer view of the distribution within languages, Figures 5-4, 5-5 and 5-6 show the percentage of questions per function for English, French and Spanish respectively *i.e.* their relative frequency.

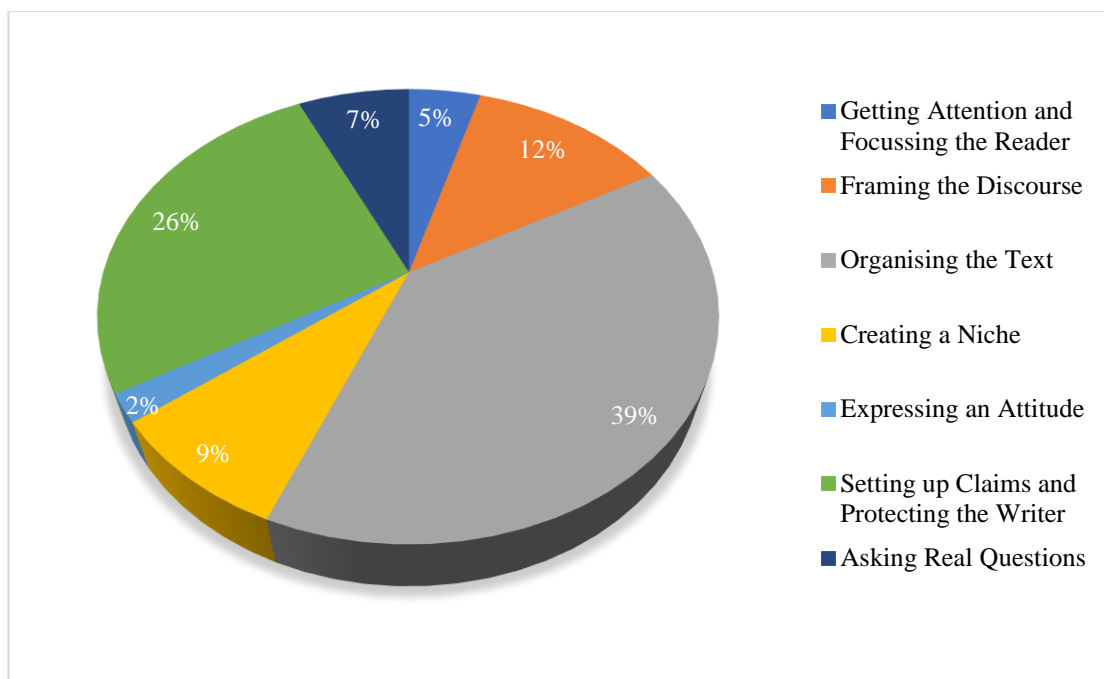


Figure 5-4 Percentage distribution question functions in engecon

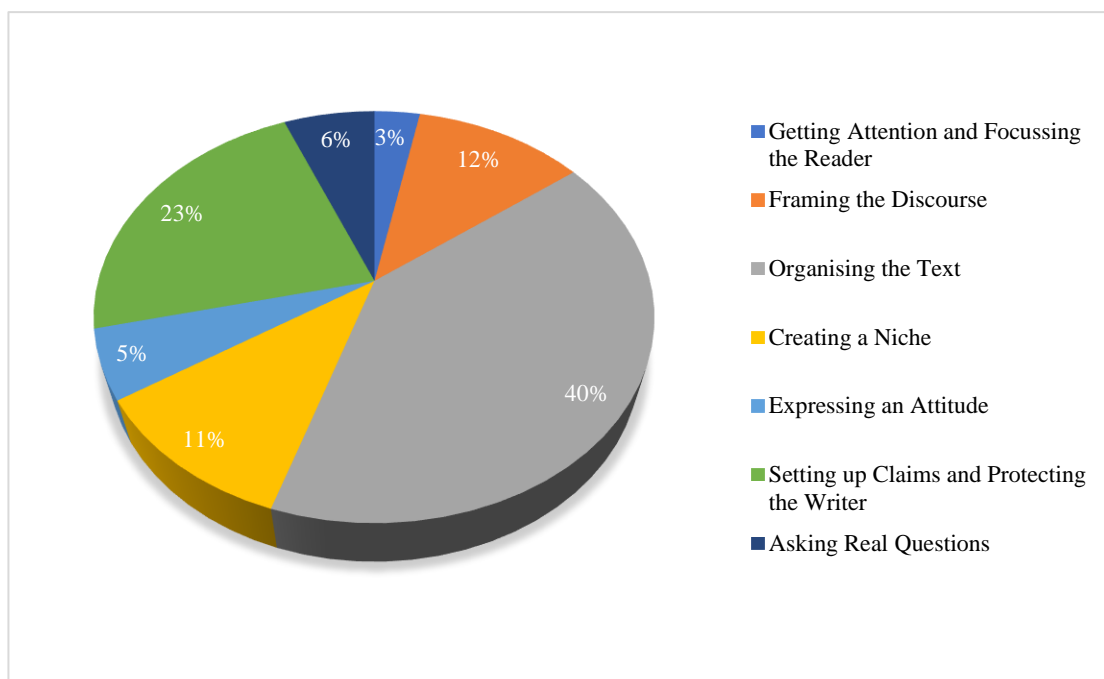


Figure 5-5 Percentage distribution question functions in frecon

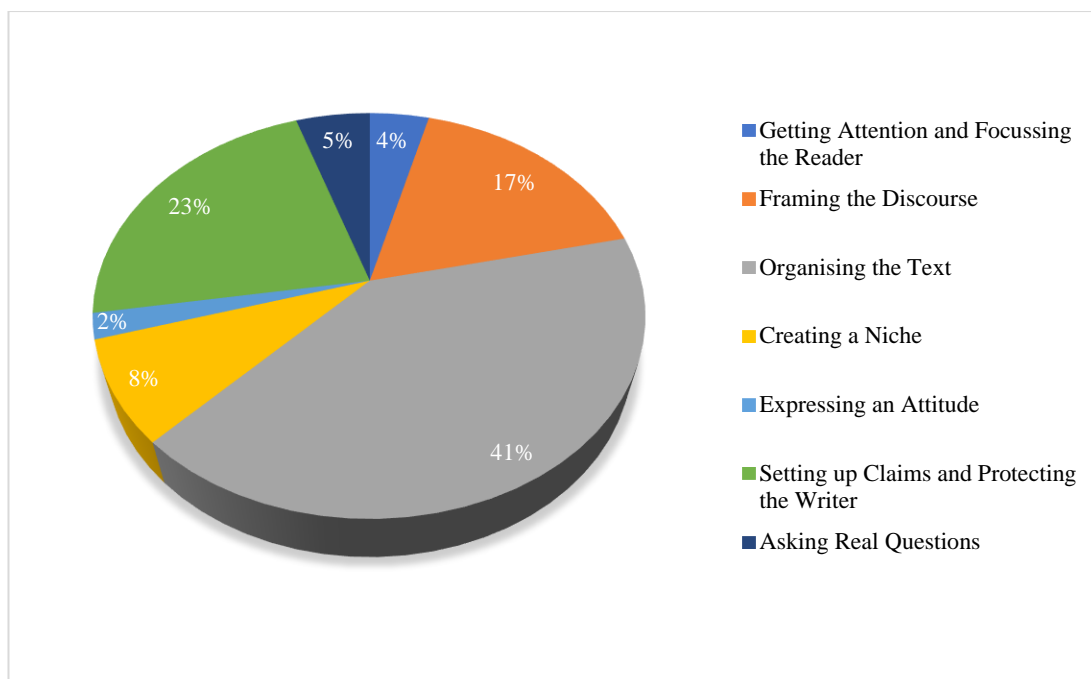


Figure 5-6 Percentage distribution question functions in Spanish

Each language uses the same five functions in descending order of frequency:

1. Organising the text
2. Setting up claims and protecting the writer
3. Framing the discourse
4. Creating a niche
5. Asking real questions

For English and Spanish, getting attention and focussing the reader is the sixth most frequent function and expressing an attitude and counter-claiming is the seventh most frequent function. However, for French these two functions are inverted with expressing an attitude and counter-claiming as the sixth most frequent function and getting attention and focussing the reader as the seventh most frequent function. However, it should be noted that the difference in frequency between those two functions in each language is very small. Comparing relative frequencies of function within and across languages, it is also clear that the proportion of questions used to get attention and focus the reader in English is higher than any other language. Furthermore, the proportion of questions used to express criticism in French is higher than any other language and the proportion of questions used to frame the discourse in Spanish is higher than any other language. Overall however, it is reasonable to argue that there is consistency in function use and this is an important finding as it has informed the *tertium comparationis* for Chapter 6.

5.4 Question sentence length

Sentence length presents some interesting variation in English, French and Spanish. The French subcorpus contains questions that are longer than those in English and Spanish which is reported in Table 5-2.

Table 5-2 Question length in words in KIAP-EEFS

Sentence length	English	French	Spanish
Average question length in words	18	23.7	21.1
Median question length in words	16	22	15
Mode question length in words	10	16	7

Questions on average are shorter in English than French and Spanish, which share a similar length. However, looking at the median of every question, which is perhaps more reliable given the existence of outlier question lengths, English and Spanish questions are closer to one another in length. The mode, which represents the most common exact length, also reveals that within English, the most common length is 10 words, in French the most common length is 16 words and in Spanish, the most common length is 7 words. However, this is potentially not so illuminating with such a small data set, as if questions differ in length by just one word, they are not captured in the mode. In a categorical view, Table 5-3 presents the amount of questions with fewer than 15 words, between 16 and 25 words or 26 words or more.

Table 5-3 Question length ranges in KIAP-EEFS in WPM and percentage of all questions

Sentence length	English	% English	French	% French	Spanish	% Spanish
15 words or fewer	249	49.3	83	25.6	126	53.1
16-25 words	138	27.4	133	41.1	56	23.45
26 words or more	118	23.3	107	33.3	56	23.45

Most questions in English and Spanish make use of fewer than 15 words, while French questions tend to be longer. In fact, 49.3% and 53.1% of all questions in English and Spanish, respectively, occur in 15 words or less while in French only 25.6% of questions are of this length. Conversely, 41.1% of questions in French are fit within the range of

16-25 words while only 27.4% and 23.45% of questions in English and Spanish, respectively, are of this length.

This difference in sentence length across languages is interesting as, while the data initially showed a convergence surrounding distribution and function, when a criterion like sentence length is considered, there is evidence of difference across languages.

5.5 Question type and form

Across the three subcorpora, questions are categorised according to two types: direct and indirect, and five forms: content, polar, declarative, alternative and elliptical questions. Each subcorpus uses mainly direct questions and the English data use more direct and indirect questions than French, while the French data use more direct and indirect questions than Spanish, according to Figure 5-7.

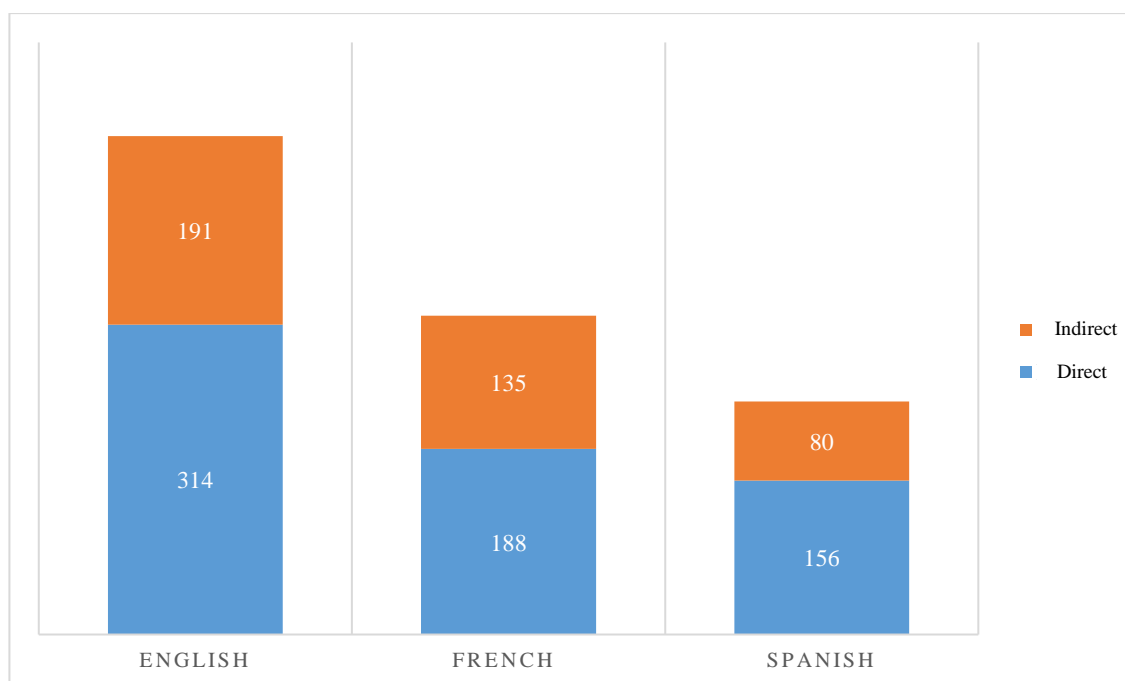


Figure 5-7 Direct and indirect questions in KIAP-EEFS in WPM

Furthermore, looking at distribution within each language, 43.2% of French, 37.8% of English, and 33.7% of Spanish questions are indirect, as indicated in Table 5-4. Overall, direct questions are more common, but a higher distribution of questions in French are indirect.

Table 5-4 Percentage direct and indirect questions in KIAP-EEFS

	English	French	Spanish
Direct questions	62.2	56.8	66.3
Indirect questions	37.8	43.2	33.7

Each language appears to have similar patterns of usage across the forms of questions used to pose direct and indirect questions, indicated in Figure 5-8. However some important difference to appear. For example, English and Spanish make more use of direct content questions while French uses mainly direct polar questions. Elliptical questions also emerge in French as more frequent.

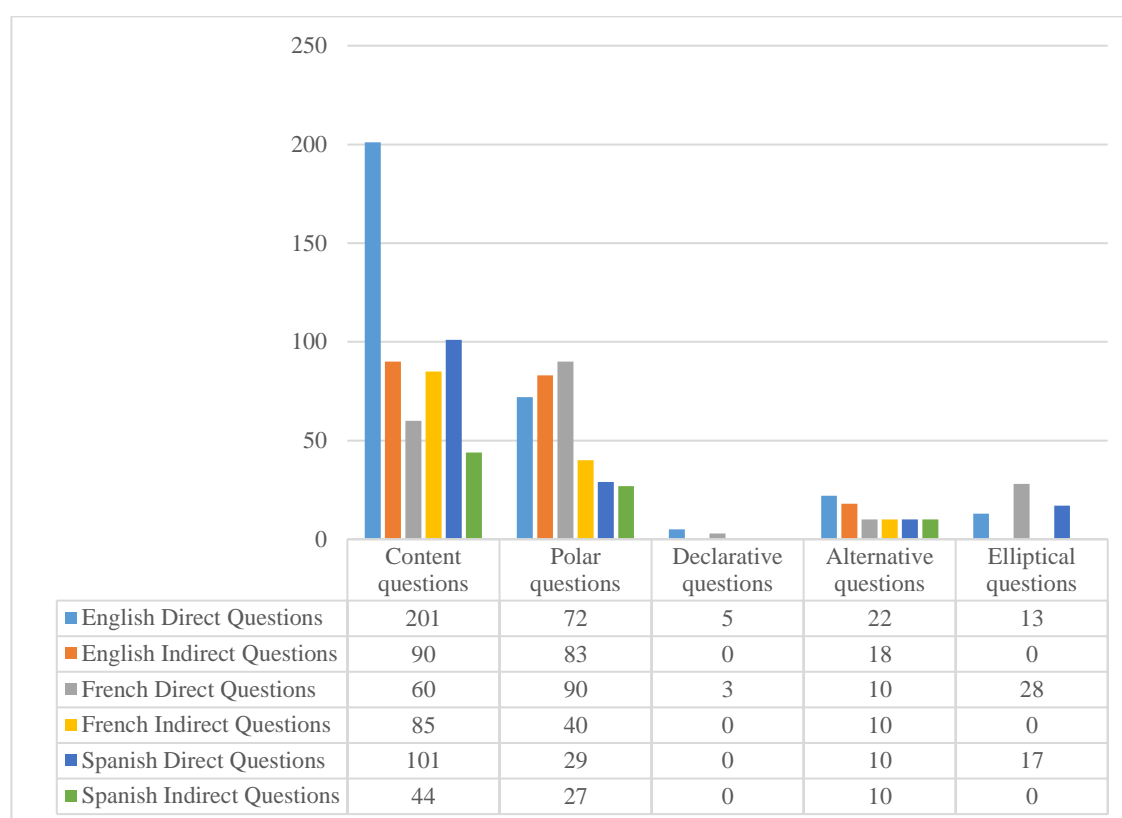


Figure 5-8 Question types in KIAP-EEFS in WPM

Comparing these languages more closely, Table 5-5 shows the percentage of direct and indirect questions in KIAP-EEFS per form in order to demonstrate how question types and forms are distributed within each language.

Table 5-5 Percentage direct and indirect questions in KIAP-EEFS per type and form

	Content	Polar	Declarative	Alternative	Elliptical
English direct questions	39.8	14.4	1	4.5	2.5
English indirect questions	17.9	16.4	0	3.5	0
French direct questions	17.8	27.9	0.8	3.1	8.5
French Indirect questions	26.4	12.4	0	3.1	0
Spanish direct questions	42.9	12.2	0	4.1	7.1
Spanish indirect questions	18.4	11.2	0	4.1	0

English and Spanish make the most use of direct content questions which account for 39.8% and 42.9% of all questions identified in each language respectively. French however, contains much fewer examples with only 17.8% of all questions in French being direct content questions. In contrast, only 17.9% and 18.4% of all questions in English and Spanish are indirect content questions, such questions in French account for 26.4% of all questions. In terms of polar questions English and Spanish follow a similar pattern with no great difference between direct and indirect polar question usage. In French 27.9% of all questions are direct polar and only 12.4% are indirect polar. Declarative questions are rarely used in general, and alternative questions are similarly infrequent for direct and indirect questions in each language. For elliptical questions there are no indirect questions and, relatively, French and Spanish make more use of elliptical questions than English. In fact, 8.5% and 7.1% of questions in French and Spanish are elliptical and only 2.5% of English questions are elliptical. Importantly, each language shares a pattern of decreasing use in types from content, polar, alternative, elliptical and declarative questions. Overall, there are both similarities and differences in the use of direct and indirect questions and different question types in each language. Therefore, once again, this finding may lead to questioning the view that questions in each language behave similarly.

5.5.1 Indirect questions: illocutionary force indicating devices

While direct questions are signalled by a question mark, indirect questions, as detailed in Section 4.2.3 are identified by illocutionary force indicating devices in the form of verbs,

nouns, and adjectives. In the case of each language, verbs are most frequent, followed by nouns and adjectives, according to Table 5-5.

Table 5-5 IFIDs that generate indirect questions in KIAP-EEFS in WPM and percentage of all questions

IFID form	English	%English	French	%French	Spanish	%Spanish
Noun	40	8	15	4.7	22	10.2
Adjective	15	3	0	0	5	1
Verb	136	26.9	117	36.4	53	22.5

5.5.2 IFID: nouns

English uses twice as many nouns to ask questions than French and Spanish. However, questions raised with nouns account for 8% of all English questions, 4.7% of all French questions, and 10.2% of all Spanish questions. Therefore, nouns emerge as somewhat important for both English and Spanish. Table 5-6 shows the most common noun forms used to create question in each language.

Table 5-6 Nouns used to create indirect questions in KIAP-EEFS in WPM

Rank	English		French		Spanish	
1	Question	35	Question	15	Cuestión	5
2	Investigation	3			Pregunta	5
3	Test	3			Proporción	2
4					Cuestiones	2

The noun “question” is the only finding of note when comparing across languages. The following examples show how “question” is used:

The first extension concerns the question of whether the proposition in Theorem 1 can be reversed. (engecon07)

La question qui se pose alors est de savoir quel sera le "prix" de cette action évalué par les différents individus (frecon08)

Ahora bien, la cuestión es cómo conseguir estos objetivos (specon10)

From these examples in English, French, and Spanish, the word question/question/cuestión is being used to set up an indirect question. These indirect questions

serve rhetorical purposes in each language. An interesting finding surrounding these cognates arises as neither English nor French use question as a verb where Spanish uses both the noun and verb parts of speech of *cuestión*.

5.5.3 IFID: adjectives

Adjectives are rarely used in KIAP-EEFS as indirect question IFIDs. There are no examples in French and few examples in English and Spanish which account for only 3% and 2% of all questions found in each language, respectively. The only examples in English is the use of “interested” and “clear” and the only example in Spanish uses *cuestionable* to create indirect questions in the following way:

It is not clear whether inflation rises or falls with the amount of time left in office.
(engecon27)

Resulta cuestionable la medida en que esta recomendación mejore efectivamente el aprovechamiento de los estudiantes (specon44)

It is important to note that due to the minimal use of adjectives to create questions in the corpus, there is little of substance to be drawn from those that do occur. That being said, an investigation of such parts of speech as question IFIDs would make for an interesting future study in a larger data set.

5.5.4 IFID: verbs

Verbs used in KIAP-EEFS to create indirect questions contain more notable patterns as can be seen in Table 5-7.

Table 5-7 Verbs used to generate indirect questions in KIAP-EEFS in WPM

Rank	English		French		Spanish	
1	Ask	25	Interroger	28	Preguntar	12
2	Examine	20	Déterminer	15	Investigar	10
3	Explore	17	Savoir	15	Precisar	7
4	See	10	Montrer	15	Evaluar	5
5	Tell	8	Demander	15	Conocer	5
6	Check	5	Tester	8	Saber	5
7	Gauge	5	Poser	5	Examinar	2

8	Test	5	Étudier	3	Cuestionar	2
9	Understand	5	Examiner	3	Ahorrar	2
10	Determine	5	Analyser	3	Determinar	2

Overall, there is higher reuse of verbs and strong similarities in the high frequency verbs used to ask indirect questions. For example, the verbs: ask/*demander/interroger/preguntar/cuestionar*, which share very similar meanings, all rank quite highly among verbs used to create indirect questions. Moreover, they work in very similar ways as can be seen here:

Finally, we briefly characterize some stylized facts regarding our estimated stocks and ask whether there are trends in net foreign assets and shifts in their composition over time. (engecon02)

Si ces résultats devaient toutefois être confirmés, il serait légitime de se demander s'ils reflètent la spécificité du dispositif français de formation continue ou si leur portée est plus générale. (frecon22)

Il nous semble donc intéressant de revenir aux fondements de la microéconomie du consommateur et d'interroger cette dernière à la lumière de la soutenabilité (frecon02)

De estas las principales son de Estados Unidos con lo cual hay que preguntarse seriamente si el proceso de globalización no debiese llevar más bien la etiqueta “Made in USA” y ser considerado simplemente como un nuevo producto que venden las transnacionales de dicho país (specon31)

El cambio tecnológico en el mundo se está traduciendo en una mayor brecha salarial entre los trabajadores más calificados y los menos calificados, lo que cuestiona la necesidad de la intervención gubernamental, tanto en los países desarrollados como en los que están en desarrollo (specon42)

Spanish also makes use of the verb *investigar* with no examples of this verb cognate being used in English or French. While this may be surprising for English, for French it is less so given that there is not cognate form of “investigate”. In the subsequent examples we can the use of this verb to create a question by expressing the need to discover something:

Primero, intentamos poner a prueba simultáneamente los pronósticos principales de dos grupos de modelos de las crisis (los llamados modelos de primera y segunda generaciones de crisis), a fin de investigar si estos modelos teóricos nos han ayudado a entender por qué ocurren las crisis y elucidar la cuestión de si estos modelos son complementos o sustitutos en la explicación de la presencia de crisis. (specon37)

French makes use of *analyser* in a similar fashion:

Nous allons analyser comment les solutions en [...] de ce système d'équations varient en fonction de N. (frecon43)

This use of a verb and a question word to reveal information is a relatively common means to pose and make use of an indirect question, as discussed in Section 2.10.2

English makes more frequent use of “examine” with French and Spanish using *examiner/ examinar* much less frequently. All three however create a question by indicating what the writers want to know through the proposed examination. The following indirect questions exemplify this.

Building on these predictions, we examine whether there is a negative relationship between the strength of FPRs and labor flows from the source country to the affiliates. (engecon04)

Au préalable, nous examinons quelles sont les questions spécifiques posées par le système de retraite en matière de redistribution (frecon15)

Examinare ahora como es afectada la tasa de interés a largo plazo por las políticas fiscales de manera congruente con el crecimiento de estado estacionario. (specon49)

English and French also bear some similarities in the use of *test/tester* which does not feature in Spanish as it is not a cognate form. Its close equivalent, *probar*, does not feature either, indicating that “test” is not a verb used to raise questions in Spanish, at least in these data. In the following examples, *test/tester* is being used to reveal information like many of the verbs already discussed here:

The final alternative specification, tests whether Fed interventions that are coordinated with one or both of the other G-3 central banks are more influential than unilateral interventions. (engecon46)

Si l'hypothèse de marchés complets est rejetée, il est donc intéressant de tester si le métayage favorise le partage des risques ou non. (frecon29)

French makes much greater use of verb *déterminer* whose cognate does also occur in the list of verbs in English and Spanish, but to a much lesser degree. In all three, however, the verb seeks to prove or show something, again related to the idea of discovery or revelation of information.

Notre objectif est maintenant de déterminer les prix d'équilibre (frecon43)

There is also evidence of verbs being used in each language that do not feature in the others. French uses high ranking verbs *savoir* and *montrer* to create indirect questions such as:

La question qui importe, et qui est celle étudiée dans l'article, est donc celle de savoir si une banque paie plus cher un besoin accru de financement externe (frecon48)

Nous montrons comment estimer les paramètres de préférences des ménages sous l'hypothèse nulle par une méthode de variables instrumentales (frecon29)

In these examples the authors are studying the way in which decisions are checked. They want to know if banks pay too much and want to show how to make certain estimates. Each one of these examples raises an indirect question in order to guide the reader's thinking.

Similarly, English features verbs “explore”, “see”, and “tell” to create indirect questions such as the following:

We explore whether FPRs play a more important role in protecting knowledge that is transferred outside the source country and firm relative to knowledge inside the source country and firm. (engecon04)

The aim of the present paper is to see whether the change of currency regime slowed the growth of the volume of Anglo–Irish trade, *ceteris paribus* (engecon28)

The data will then tell us whether, within the theoretical framework developed, these proxies for lobbying costs influence protection. (engecon40)

Again, each verb is used to express an intention to discover the answer to something. This evokes a question in the reader's mind, helping guide their thinking.

Finally, Spanish sees the use of verbs like *precisar* to create questions. In the example that follows, the authors want to “specify whether tourist activity is really profitable”. That is to say, they want to answer the question “is tourist activity really profitable?”

Es indudable que las dificultades estadísticas que habrá que superar, a fin de precisar si la actividad turística es realmente rentable y si ésta es preferible a las distintas alternativas existentes, serán, frecuentemente, considerable (specon11)

Overall, there is evidence of shared use of cognate and corresponding verbs to raise questions in English, French and Spanish. However, there is also evidence of verbs that are only used in certain languages. Where this occurs, it tends to occur in high frequency and/or rank.

5.6 Location in the text

When considering the location of questions, each subcorpus sees most questions occurring in the middle sections of the research article, rather than the title, introduction or conclusion, as Table 5-8 presents.

Table 5-8 Questions per location in KIAP-EEFS in WPM and percentage of all questions

Section	English	%English	French	%French	Spanish	%Spanish
Title	23	4.5	10	3.1	10	4.1
Introduction	103	20.4	95	29.5	70	29.6
Middle	334	66.1	198	61.2	153	64.3
Conclusion	45	9	20	6.2	5	2

In fact, 66.1%, 61.2% and 64.3% of all questions in English, French, and Spanish, respectively, occur within the middle sections. However, there remain some useful differences. Interestingly while more examples of questions do occur in introduction sections in English than any other language, proportionally, French and Spanish see 29.5% and 29.6% of all questions occur in introductions with only 20.4% in English. For titles, English and Spanish see 4.5% and 4.1% of all questions occur therein, while French sees only 3.1% of questions in titles. Finally, another the key differences occurs in questions used in conclusions which account of 9% of English, 6.2% of French, and only 2% of Spanish questions identified. Of course, given that only 78% of all texts in specon contain conclusions this is not surprising. This contrasts directly to engecon and frecon in which 98% and 86% of texts contain explicit conclusion sections. Overall, this section shows that there are differences in the use of questions across languages with similarities and differences occurring between languages.

5.7 Passivity, tense and verbal modality, and sentence type

This section considers the passivity, tense and verbal modality, and sentence type of all questions in English, French and Spanish in order to identify any patterns. Overall, there is great convergence in patterns across the subcorpora with less passivity, more present tense, and more simple and complex question sentence structures. There is however, some divergence from this pattern.

5.7.1 Passivity

When referring to passivity, there are three potential categories in the data: 1) not passive; 2) passive and 3) unclear. Passive means that there is either a “be + past participle” structure or a *se*-, reflexive, or *on*-passive and not passive means that there is an absence of these structures. Questions categorised as “unclear” are direct questions or IFIDs that

omit tense and/or verb phrases in the question generating word or phrase. The distribution of passivity across the three languages is largely the same, as Figure 5-12 shows.

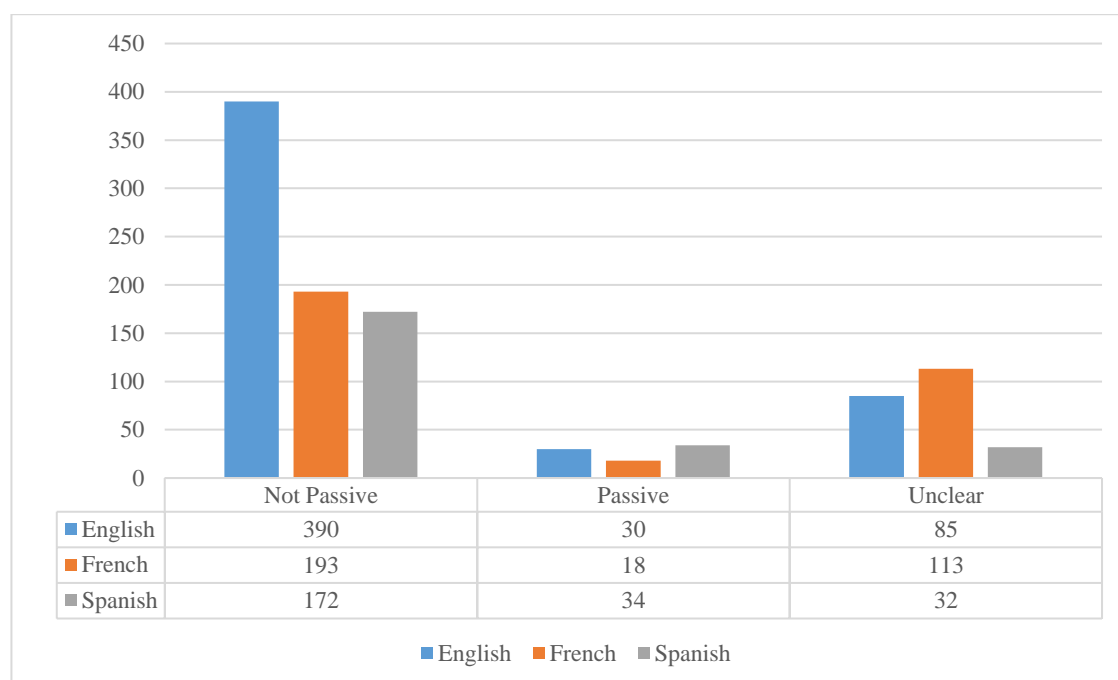


Figure 5-9 Passivity in questions in KIAP-EEFS in WPM

There is a clear absence of passivity, where each subcorpus makes little use of passivity in questions. In fact, only 6%, 5.4% and 14.3% of all questions in English, French, and Spanish contain passivity, as presented in Table 5-9.

Table 5-9 Percentage passivity in KIAP-EEFS

KWIC type	English	French	Spanish
Not passive	77.1	59.7	72.4
Passive	6	5.4	14.3
Unclear	16.9	34.9	13.3

Spanish use more passivity proportionally and French uses more “unclear” than any other language. In the case of the passivity, for all three languages, it is very infrequently used to raise questions. Where it is used, questions take different forms in different languages:

How are our SBTC results affected by not examining factor-supply changes?
(engecon23)

La rémunération du directeur doit-elle être considérée comme un salaire ou comme une part du profit ? (frecon18)

On s’interroge alors sur l’existence et les caractéristiques de pénalités permettant d’améliorer la situation (frecon30)

¿Por qué, independientemente de la nueva legislación en la materia, los sistemas de tratamiento no han sido adoptados aun cuando pueden redituvar ganancias? (specon18)

¿Qué consecuencias se derivan desde el punto de vista del razonamiento lógico para cada uno de los modelos? (specon25)

In each language, the examples show passivity using a structure of “verb to be” and “past participle” *i.e.* in Spanish there is: *no han sido adoptados*. Moreover, in Spanish and French, there is evidence of the passivity through a reflexive structure and the also in the use of *on* in French which obfuscate the agency. For unclear passivity, the following example shows how neither active nor passive agency is determinable:

Le consommateur, producteur de soutenabilité? (frecon02)

There is no verb or tense to indicate passivity in the above example. As such, it cannot be categorised. Overall, there is a general convergence toward passivity not being used in questions.

5.7.2 Tense and verbal modality

When considering tense, there is evidence of questions that use:

- past tense;
- present tense;
- future tense;
- no tense;
- and verbal modality.

While most questions converged towards the use of present tense in KIAP-EEFS, there is evidence of questions using different tenses in different languages, presented in Table 5-10 and Figure 5-13.

Table 5-10 Percentage tenses in KIAP-EEFS

KWIC type	English	French	Spanish
Past	4.5	4.7	3.1
Present	58.7	37.9	42.9
Future	3.5	0.8	2

No tense	17.9	35.6	38.8
Modality	15.4	21	13.2

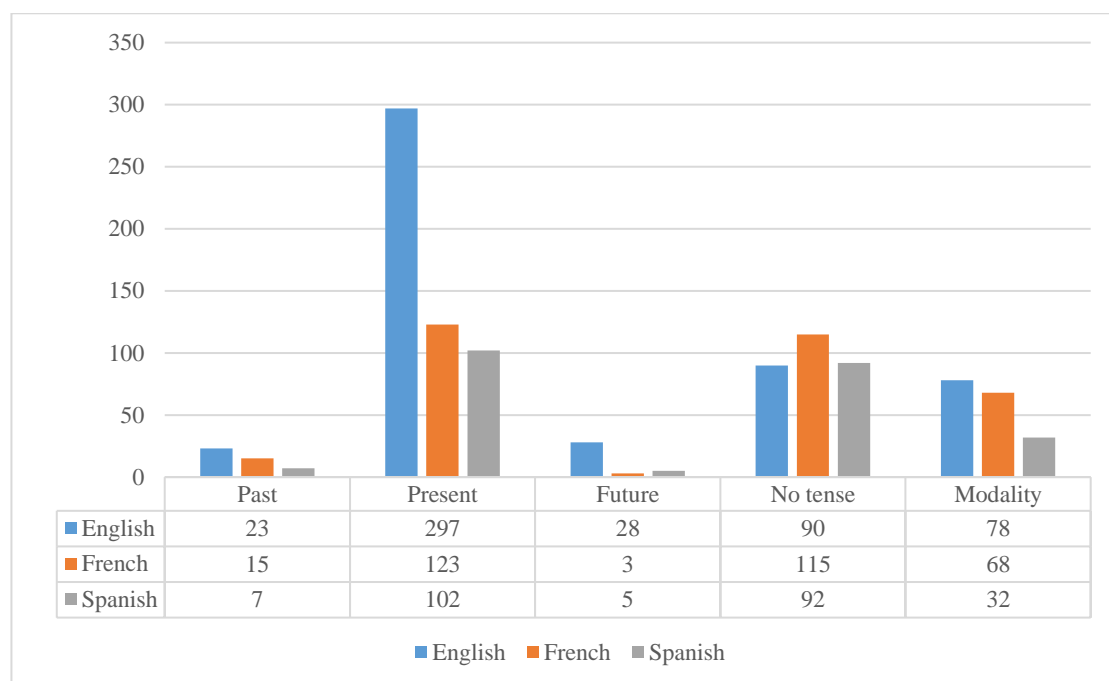


Figure 5-10 Tenses in KIAP-EEFS in WPM

Most questions in the future occur in English, which account for only 3.5% of all questions in English, *e.g.*:

The data will then tell us whether, within the theoretical framework developed, these proxies for lobbying costs influence protection (engecon40)

These are less evident in French and Spanish. Future questions account only for 0.8% and 2% of all questions in French and Spanish, respectively. Instead in French and Spanish there is evidence of many more tense-less questions as in the example presented previously:

Le consommateur, producteur de soutenabilité ? (frecon02)

These slight variations notwithstanding, each language generally acts similarly with questions being asked in the present tense and using modality where 58.7%, 37.9% and 42.9% of all questions in English, French, and Spanish respectively pose questions in the present and 15.4%, 21% and 13.2% of all questions in the respective languages ask questions using some form of verbal modality.

5.7.3 Sentence type

There are 5 categories of question type based on the clausal construction of the question. These are: 1) simple; 2) complex; 3) compound; 4) complex-compound; 5) no clause. Most questions in KIAP-EEFS were mostly either simple or complex, as can be seen in Table 5-11 and Figure 5-14.

Table 5-11 Percentage sentence type in KIAP-EEFS

KWIC type	English	French	Spanish
Simple	33.3	22.5	40.8
Complex	37.8	55.8	24.5
Compound	11.5	7	10.2
Compound-complex	13.4	5.4	20.4
No clause	4	9.3	4.1

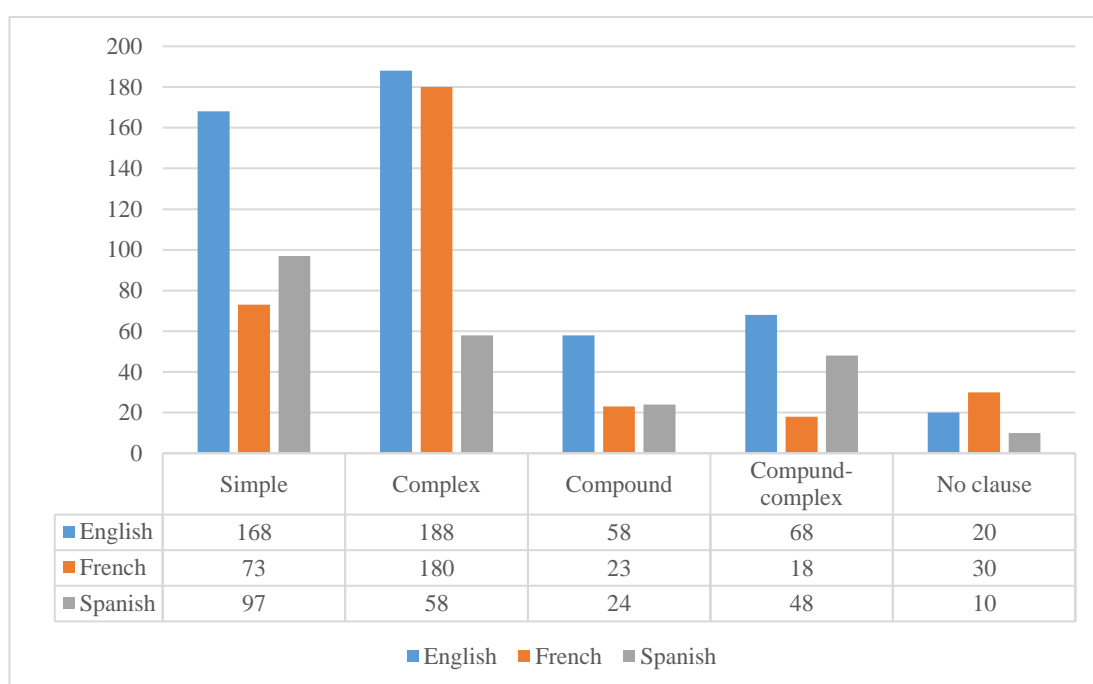


Figure 5-14 Sentence types for questions in KIAP-EEFS in WPM

55.8% of all questions in French occur as complex questions. This is quite a contrast to 37.8% in English, and 24.5% in Spanish. Such questions often contain main and subordinating clauses such as:

Au préalable, nous examinons quelles sont les questions spécifiques posées par le système de retraite en matière de redistribution (frecon15)

In fact, so many questions in French make use of this sentence structure that all the others are lower than every other language except for those with no complete clauses, such as:

Quel arbitrage efficacité-équité dans le financement et la gestion de la filière de soins?
(frecon03)

In the example question, there is no verb to create main or subordinating/coordinating clauses and this sentence type occurs most often in French. Spanish uses many more simple sentences and compound-complex sentences which account for 40.8% and 20.4% of all questions in Spanish. English relates to both languages with main and similar usage of simple and complex sentences which account for 33.3% and 37.8% of all questions in English.

5.7.4 Passivity, tense and verbal modality, and sentence type: summary

Overall, in this section there is clear evidence of similarities across languages. Little passivity is used and present tense features most highly among all questions in KIAP-EEFS. For the most part, questions are either simple or complex in KIAP-EEFS in terms of clausal construction. Nonetheless, there are interesting differences to be seen. For example, French shows most variation with questions with unclear passivity, and no tense or clause elements. Regarding sentence type French uses more complex sentences for questions than any other language and more than any other sentence type.

5.8 Manifest similarities of questions based on degree

This section reviews the key findings of this first phase of the study which outlines the manifest similarities *i.e.* the ways in which questions correspond to one another in KIAP-EEFS. From these findings it is possible to propose a maximally similar correspondence.

1. Questions occur in each subcorpus and there are enough examples and a sufficiently even distribution to render them an interesting feature of academic discourse represented within KIAP-EEFS. That being said, when compared to Spanish, questions in English are significantly more frequent.
2. Questions that occurred in each subcorpus of KIAP-EEFS use the same functions across languages with a largely similar functional distribution. However, their proportion of use does differ according to relative frequency. For example:

- a. All three languages mainly use questions to organise the text, set up claims and protect the writer, and frame the discourse in descending order frequency.
 - b. The proportion of questions used to get attention and focus the reader in English is higher than any other language.
 - c. The proportion of questions used to express criticism in French is higher than any other language.
 - d. The proportion of questions used to frame the discourse in Spanish is higher than any other language.
3. Questions in English and Spanish are a similar length and tend to be shorter than questions in French.
4. There are variations in the types of questions evident in each language. There are too many differences to list here but some examples are:
 - a. All languages use more direct questions than indirect questions. However, relatively, French uses most indirect questions where 43.2% of all questions in French are indirect, as opposed to 37.8% in English and 33.7% in Spanish.
 - b. Of direct questions, English and Spanish use mainly content questions while French uses mainly polar questions.
 - c. Of indirect questions, all three languages use mainly content questions.
 - d. There is greater use of direct content questions in English and Spanish and indirect content questions in French.
 - e. There is greater use of indirect polar questions in English and direct polar questions in French. Spanish uses a similar amount of direct and indirect polar questions.
 - f. There are no examples of declarative questions in Spanish and few examples in English and French.
 - g. For IFIDs of indirect questions, in each language verbs are the most common IFIDs, followed by nouns. Adjectival IFIDs are negligible in general. The main verbs used to create indirect questions are “ask”, *interroger*, *preguntar*, “examine”, *déterminer*, *investigar*, and “explore”, *savoir* and, *precisar*. For nouns the focus is on “question” as a noun in each language where Spanish is the only language to use *cuestión* as a verb.
5. Regarding location:
 - a. Most questions occur within the thematic middle sections in KIAP-EEFS, followed by introduction, and then the title and conclusion. This finding can be explained by the fact that opportunity to ask questions is greater in the middle sections.

- b. Proportionally, French and Spanish use more questions in introduction sections than English.
6. There is evidence of very little passivity, mostly present tense usage as well as simple and complex question sentence structures in KIAP-EEFS. This trend is consistent across the subcorpora. Some variation occurs in French with questions that have no voice, no tense, or formal complete clauses.

5.9 Identifying correspondence in question use as reader engagement

This section proposes how questions most likely correspond in English, French, and Spanish, based on KIAP-EEFS. The following brief paragraph presents what is maximally similar based on degree of similarity in questions in KIAP-EEFS.

Where English, French, and Spanish economics research articles contain questions used as a rhetorical feature of reader engagement they are primarily used to organise the text in each language. These questions are short in English and Spanish and longer in French. In terms of question type and form, each language uses direct questions with English and Spanish using content questions, and French using polar questions. Most often each language uses interrogative verbs as question IFIDs (ask, *interroger*, *preguntar*). These questions occur in the middle sections of the research paper for each language, do not use passivity in any language, use the present tense in each language, and take a complex sentence structure in English and French and a simple sentence structure in Spanish.

Of course, this is a generalisation as these are the most frequent comparable characteristics of questions identified in each language in this study. Therefore, frequency is used as the determining factor to make this decision. While it is clear that contrastive analysis' claims must be tentative and other factors could be important in making decisions about correspondence, based on this analysis, it is reasonable to argue that the nature of questions described in the above paragraph is an effective description of how questions as reader engagement correspond across English, French, and Spanish.

5.10 Conclusion

This chapter has presented the results of the first contrastive analysis of questions in KIAP-EEFS for which the analysis is outlined in Sections 5.2-5.7. This section centred on the *tertium comparationis* identified, which was the assumption that English, French, and Spanish economics research articles used questions as a rhetorical feature of reader engagement. Questions do appear to be a rhetorical feature of the three subcorpora of

KIAP-EEFS and the analysis revealed many similarities and differences in the presence of questions. These are summarised in Section 5.8 in which the manifest similarities in each language are presented. Section 5.9 builds on this analysis and presents a maximally similar correspondence for this *tertium comparationis* based on degree of correspondence *i.e.* frequency. However, while this chapter offers a clear picture of questions in economics academic writing in English, French, and Spanish, its breadth hides variations that occur across functions and within languages. As such, the next chapter looks deeper into question function by testing the *tertium comparationis* of shared question function in KIAP-EEFS.

Chapter Six

Corpus-based contrastive analysis of
question function as a shared rhetorical
feature of reader engagement in
English, French, and Spanish

6 Corpus-based contrastive analysis of question function as a shared rhetorical feature of reader engagement in English, French, and Spanish

6.1 Introduction

This chapter presents the findings of the second of two corpus-based contrastive analyses of questions in economics research articles in English, French, and Spanish. Based on the findings presented in Section 5.3, the *tertium comparationis* identified in this chapter is the assumption that English, French, and Spanish economics research articles share a functional use of questions as reader engagement. As already evidenced in Section 5.3, each language follows a similar distribution of question functions which is demonstrated in Figure 6-1.

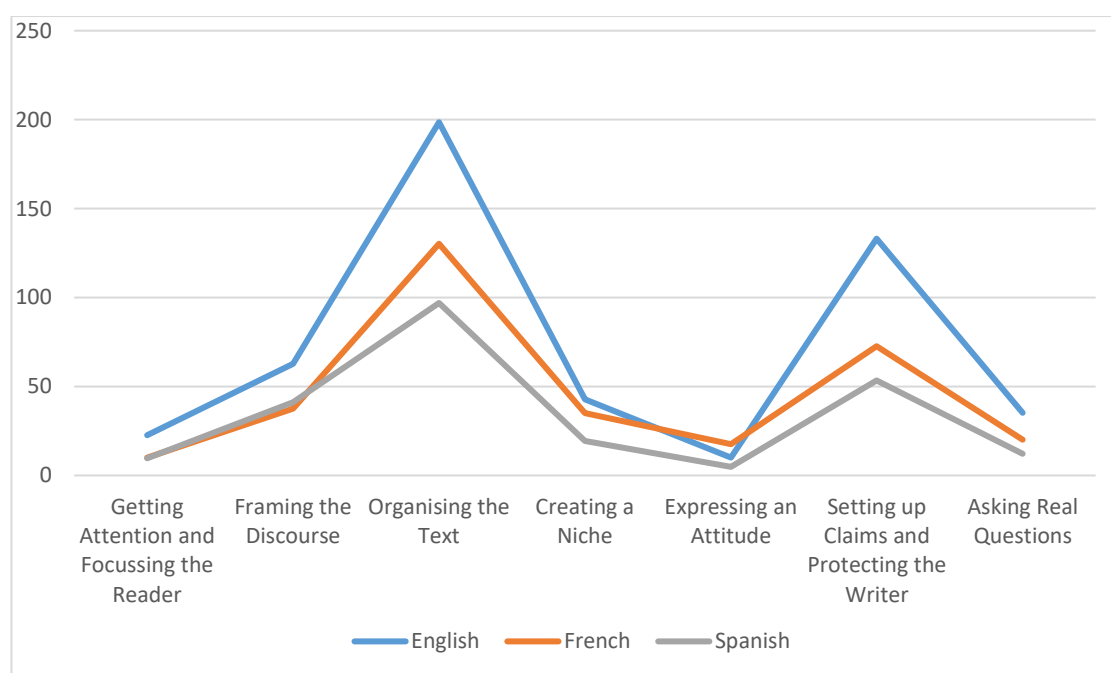


Figure 6-1 Distribution of questions per function in KIAP-EEFS in WPM

The goal of this chapter is to test this *tertium comparationis* in order to gain a deeper insight into the description of question functions and the similarities and difference in question functions across languages. Sections 6.2-6.8 present the contrastive analysis of each of the seven functions of questions. Each function is analysed in terms of frequency, length, type and form, location, passivity, tense and verbal modality, and sentence type.

Each question function is analysed with the goal to identify both the frequently used features of each function and the features of each function that are most salient. Saliency is understood as unusually frequent and infrequent uses when compared to overall distributions of questions per function within the data. In each section, a conclusion is offered to summarise the findings and identify the manifest similarities in question functions in English, French, and Spanish. Following this presentation of results, Section 6.9 offers a brief chapter conclusion.

6.2 Getting attention and focussing the reader

6.2.1 Getting attention and focussing the reader: frequency

There are more than twice as many examples of questions used to get attention and focus the reader in English than French and Spanish, as reported to Table 6-1. Overall this function accounts for 4.5%, 3.1%, and 4% of questions in English, French, and Spanish respectively. Given that this function is constrained to location, location is not a finding discussed here as it would be a case of methodological circularity.

Table 6-1 Questions to get attention and focus the reader in WPM and percentage of all questions in KIAP-EEFS

	English	French	Spanish
Total in WPM	23	10	10
Percentage %	4.5	3.1	4

6.2.2 Getting attention and focussing the reader: sentence length

Sentence length for this function is illustrated in Tables 6-2 and 6-3. Although the data set is quite small, there remain some interesting patterns in the length of questions in titles used to get attention and focus the reader.

Table 6-2 Getting attention and focussing the reader question length statistics in KIAP-EEFS

Sentence length	English	French	Spanish
Average question length in words	9.888889	16.5	8.25
Median question length in words	9	16.5	7.5

Mode question length in words	12	NA	NA
-------------------------------	----	----	----

Table 6-3 Getting attention and focussing the reader question length in WPM and percentage of all questions in KIAP-EEFS

Sentence length	English	%English	French	%French	Spanish	%Spanish
15 words or fewer	23	9.1	5	6.1	10	7.7
16-25 words	0	0	5	3.8	0	0

According to these data, questions used to get attention and focus the reader in English and Spanish hold a similar average and median word length with all questions occurring in 15 words or fewer. In fact, some quite short sentences occur in this function, such as:

How do foreign patent rights affect U.S. exports, affiliate sales, and licenses?
(engecon04)

¿Es rentable para el sector público subsidiar la educación en México? (specon24)

French however, contrasting this, has longer sentences with an average and median sitting among medium length questions.

La théorie des anticipations de la structure par terme permet-elle de rendre compte de l'évolution des taux d'intérêt sur eurodevise ? (frecon22)

Each language uses an unusually higher number of shorter questions in titles. This arises from the fact that 9.1%, 6.1%, and 7.7% of all questions shorter than 15 words in English, French, and Spanish respectively, occur in this function. Given that this function accounts for less than 5% of all questions in each language, this function contains a high proportion of short questions.

6.2.3 Getting attention and focussing the reader: question type and form

In terms of question types, each language uses only direct questions, as seen in Table 6-4. All questions in this function are direct, likely owing to their location in titles. For that reason, direct questions are unusually frequent here where this function accounts for 7.2%, 5.3% and 6.3% of all direct questions in English, French, and Spanish.

Table 6-4 Percentage direct questions in getting attention and focussing the reader in KIAP-EEFS

Question category	%English	%French	%Spanish
-------------------	----------	---------	----------

Direct questions	7.2	5.3	6.2
------------------	-----	-----	-----

Regarding question forms, there are interesting findings worth discussing, which are reported in Table 6-5 and Figure 6-2

Table 6-5 Percentage direct questions in getting attention and focussing the reader per type and form in KIAP-EEFS

	Content	Polar	Declarative	Elliptical
%English direct questions	3.8	10.3	50	40
%French direct questions	0	8.3	0	9.1
%Spanish direct questions	2.4	8.3	0	28.6

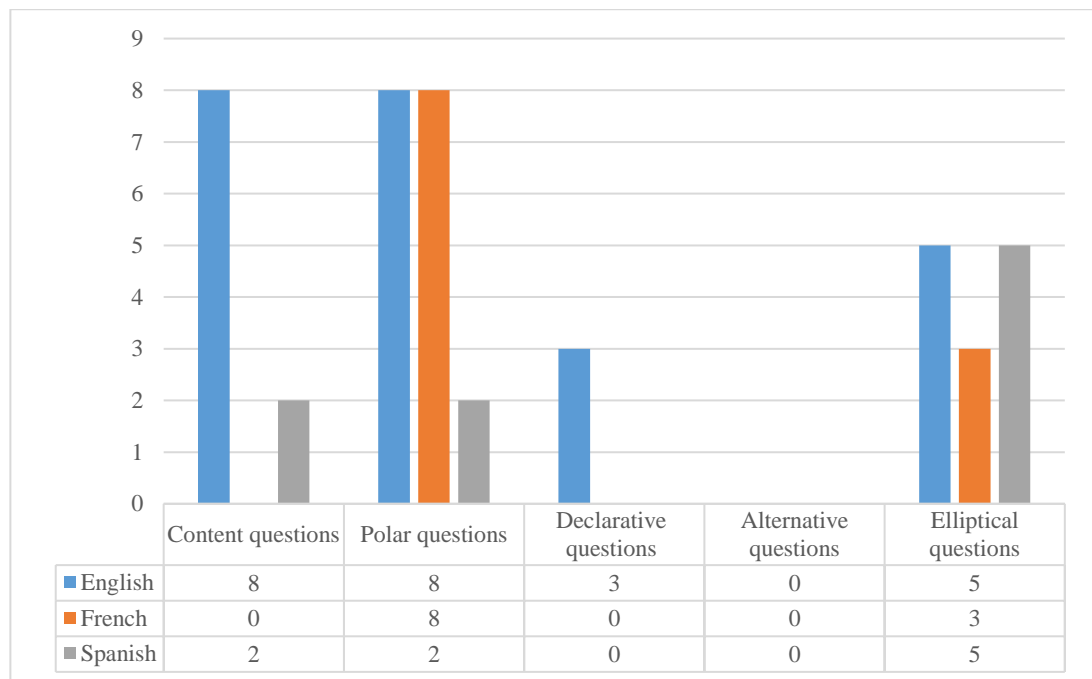


Figure 6-2 Question types and forms in getting attention and focussing the reader in KIAP-EEFS in WPM

While English uses both direct content and polar questions, this function has a more noteworthy relationship with polar questions. That is because this function accounts for only 3.8% of all direct content questions, but 10.3% of all direct polar questions in English. In this function, it is of note that 40% of the elliptical questions in the English subcorpus are found in this function. As such, elliptical questions emerge as a strong generic feature of titles. Spanish mirrors English where this function accounts for only

2.4% of all content questions, 8.3% of all polar questions, and 28.6% of the elliptical questions in the Spanish subcorpus. However, while this function also accounts for 8.3% of polar questions in French, it does not include any content questions and does not share this higher frequency of elliptical questions as only 9.1% of all elliptical questions are found in this function. This is because the remaining 90.9% elliptical questions occur in questions that organise the text in French. That being said, elliptical questions are still unusually frequent here, considering the amount of questions that get attention and focus the reader in general. Overall, the unusually frequent use of polar and elliptical questions makes for an interesting finding and it would be worth considering these question forms in a larger corpus of research article titles to see if this pattern holds.

To gain a clearer view of how elliptical questions are used in each language, here are some examples:

Macroeconomic fluctuations: Demand or supply, permanent or temporary?
(engecon22)

Projection de trajectoires économiques par microsimulation. Quelle équité pour les retraites? (frecon15)

La promoción de la actividad turística: ¿una política de desarrollo acertada? (specon11)

In English and Spanish examples, the title begins with a statement, followed by a colon and then an elliptical question. In French, the title behaves similarly, though with a period and not a colon. Beyond elliptical questions, polar questions, as in the following examples, are an evident feature of this function where they are serving to focus the readers' attention directly on the question the paper addresses:

Is it possible to redistribute the gains from trade using income taxation?

Les contrats à durée déterminée et les coûts de licenciement nuisent-ils à l'embauche stable ? (frecon07)

6.2.4 Getting attention and focussing the reader: passivity, tense and verbal modality, and sentence type

6.2.4.1 Passivity

Questions used to get attention and focus the reader in KIAP-EEFS do not show passivity in any language with the majority being either not passive or unclear. Some examples follow:

Macroeconomic fluctuations: Demand or supply, permanent or temporary? (unclear; engecon22)

La promoción de la actividad turística: ¿una política de desarrollo acertada? (unclear; specon11)

Les contrats à durée déterminée et les coûts de licenciement nuisent-ils à l'embauche stable ? (not passive; frecon07)

Owing to the prevalence of elliptical questions in English and Spanish, this function accounts for 5.9% and 15.4% of all questions in English and Spanish that contain what is categorised here as unclear passivity, as presented in Table 6-6. In contrast, this is unusually infrequent in French, where this function only captures 2.2% of unclear passivity.

Table 6-6 Passivity in getting attention and focussing the reader in WPM and percentage of all questions in KIAP-EEFS

Passivity	English	%English	French	%French	Spanish	%Spanish
Not passive	18	4.5	8	3.9	5	2.8
Passive	0	0	0	0	0	0
Unclear	5	5.9	3	2.2	5	15.4

6.2.4.2 Tense and verbal modality and sentence type

In terms of tense and modality, all questions with a clear tense inflection in this function use the present tense. This is consistent with overall patterns of use identified in Section 5.7.2. Finally, in terms of sentence type, there is no clear pattern which occurs across languages, as Table 6-7 and Figure 6-3 show.

Table 6-7 Percentage sentence type in getting attention and focussing the reader in KIAP-EEFS

Sentence type	English	French	Spanish
%Simple	10.4	0	5
%Complex	0	2.8	4.2
%Compound	0	11.1	10
%Compound-complex	0	0	0
%No clause	25	8.3	0

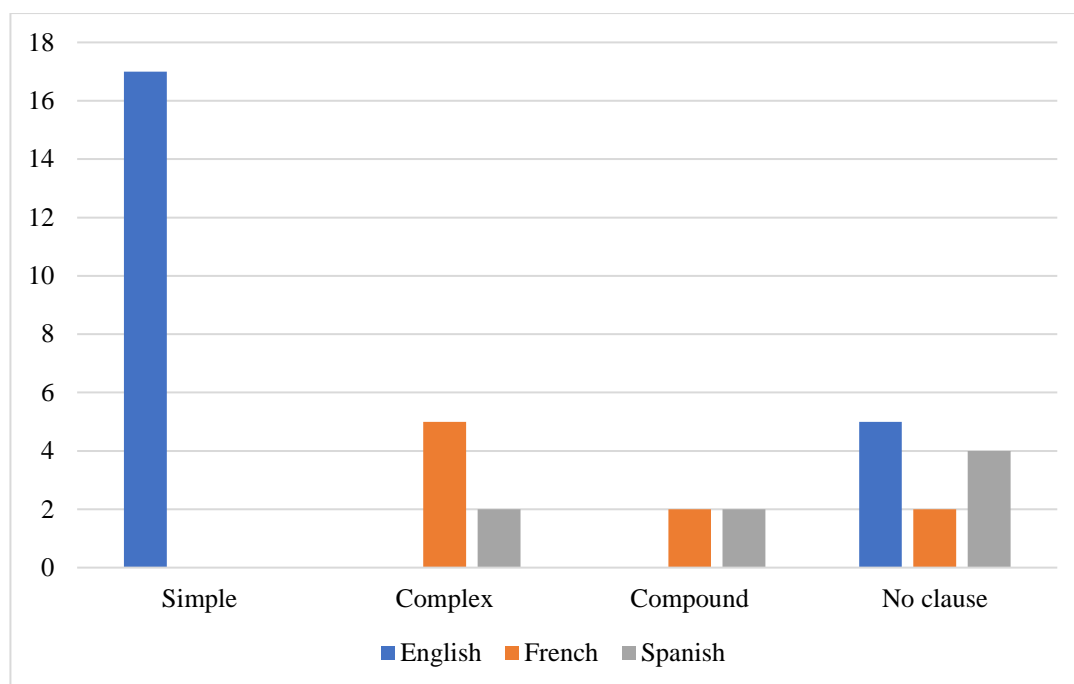


Figure 6-3 Sentence type in getting attention and focussing the reader in KIAP-EEFS in WPM

Overall, English uses mainly simple sentences. In fact, 10.4% of all questions with simple sentences in English occur in this function. In contrast both French and Spanish make proportionally greater use of questions with compound sentence types. As such, 11.1% and 10% of all compound sentence types in French and Spanish occur in this function. Otherwise, questions with sentences types with no clauses are worth mentioning: 25% and 8.3% of these in English and French occur in this function, with no examples in Spanish. An English example with a simple sentence structure is as follows:

Is it possible to redistribute the gains from trade using income taxation? (engecon05)

French and Spanish contain compound sentences, as in this Spanish example:

¿Dónde trabajaban hombres y mujeres a principios de los noventa, y dónde seis años después? segregación ocupacional por género, en México de 1991 a 1997 (specon27)

In the previous example, the use of y (and) is coordinating.

Overall, quite notable differences from the overall trends are seen here in this function, regarding sentence type. This likely due to the use of elliptical questions which increase the number of clauseless sentences in English and French. Also, noteworthy is the lack of simple sentences in French and its use of compound sentences where, unlike the English, French does not seem to make less complicated title questions.

6.2.5 Getting attention and focussing the reader: conclusion

Based on frequency, the following items capture the corresponding manifest similarities for this function across languages:

1. Questions in this function are shorter when compared to other functions. Moreover, they are shorter in English and Spanish when compared to French.
2. All languages use direct questions.
3. English uses equal amounts of direct polar and direct content questions, French uses direct polar questions, and English and Spanish use elliptical questions.
4. There is no evidence of passivity in any language.
5. All languages use mainly present tense.
6. English uses mainly simple sentences, French uses mainly complex sentences and Spanish uses an equal amount of complex and compound sentence types.

In terms of unusually frequent and infrequent features, the following distinctive characteristics emerge for this function when compared to other functions:

1. Questions are unusually short in all three languages.
2. Polar and elliptical questions are unusually frequent in English, French, and Spanish.
3. Content questions are unusually infrequent in English and Spanish.
4. Present tense is unusually frequent for all three languages.
5. English uses an unusual amount of simple sentences, while French and Spanish make unusually frequent use of compound sentences.

6.3 Framing the discourse

6.3.1 Framing the discourse: frequency

The frequency of this function is notably higher than that of questions used to get attention and focus the reader. English has more examples of this function, followed by Spanish and French. Overall, this function accounts for 17.3 % of Spanish questions, which is notably more than English and French which are closer to 12%, as reported in Table 6-8.

Table 6-8 Questions that frame the discourse in WPM and percentage of all questions in KIAP-EEFS

	English	French	Spanish

Total in WPM	63	38	41
Percentage %	12.4	11.6	17.3

6.3.2 Framing the discourse: sentence length

In terms of sentence length, questions used to frame the discourse exhibit some notable differences compared to the overall sentence length discussed in Section 5.4.

Table 6-9 Framing the discourse question length statistics in KIAP-EEFS

Sentence length	English	French	Spanish
Average question length in words	23.96	24.67	22.47
Median question length in words	26	25	15
Mode question length in words	32	30	13

Table 6-10 Framing the discourse question length in WPM and percentage of all questions in KIAP-EEFS

Sentence length	English	%English	French	%French	Spanish	%Spanish
15 words or fewer	13	5.1	10	12.1	22	17.3
16-25 words	20	14.5	13	9.4	10	17.4
26 words or more	30	25.5	15	14	10	17.4

For example, as reported in Section 5.4, English in general tends to have shorter questions in KIAP-EEFS, French has longer questions, and Spanish sits in between. However, Tables 6-9 and 6-10 show that within this function, the average and median question length in English is higher. French questions appear to be comparably as long as English questions that frame the discourse. This corresponds with the overall pattern of question length in French presented in Section 5.4. Spanish questions in this function tend to be shorter. Looking at the consistency in average, mode, and median length in Table 6-9, it is also clear that there are no outliers noticeably skewing this finding, in English and French. That being said, French appears to have fewer questions with 26 words or more than English, reported in Table 6-10. For the average and median to be so high in French,

this means that those questions of 26 words or more tend have much more than 26 words. For Spanish, it is worth noting that outliers skew the average length, as can be seen in Table 6-9. The median of 15 words matches the finding that most Spanish questions in this function tend to be composed of fewer than 15 words. This is consistent with the overall length of questions presented in Section 5.4.

Although this function accounts for 12.4% of all questions in English, it only contains 5.1% of questions which use 15 words or fewer, while it uses 14.5% and 25.5% of questions with 16-25 words and 26 words or more, respectively. Therefore, for this function in English, questions are unusually long. For French, this function accounts for 12.1% of questions which use 15 words or fewer and 14% of questions which use 26 words or more. This is especially noteworthy given this function only accounts for 11.6% of all questions in French. Finally, in Spanish, while these questions tend to be shorter, they are not unusually short. For Spanish, this function accounts for 17.3% of all questions in Spanish and includes 17.3% of all questions of 15 words or fewer and 17.4% of both 16-25 and 26 or more question lengths. This implies that this function does not use question lengths that diverge from general patterns of use.

6.3.3 Framing the discourse: question type and form

With regard to question types, there are interesting similarities and differences to be drawn from Table 6-11 and Figure 6-4. Overall, these data present a very clear distinction with English using mainly indirect questions and French and Spanish using mainly direct.

Table 6-11 Percentage direct and indirect questions in framing the discourse in KIAP-EEFS

Question category	English	French	Spanish
%Direct questions	4.8	16	20
%Indirect questions	25	7.4	12.1

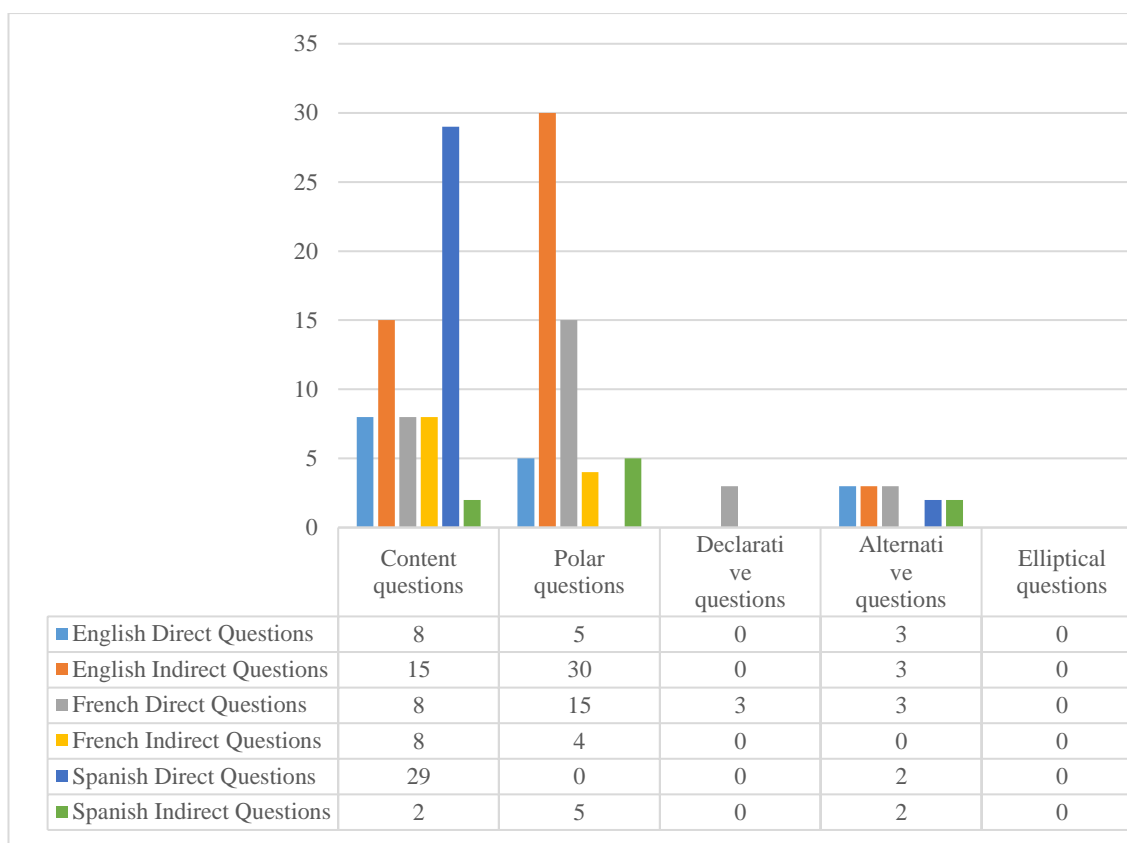


Figure 6-4 Question types and forms in framing the discourse in KIAP-EEFS in WPM

In English, indirect questions are particularly noteworthy. While for French, direct polar questions, and for Spanish direct content questions are noteworthy. For English, indirect polar questions in this function account for 36.9% of all indirect polar questions in the entire English subcorpus, as presented in Table 6-12. Furthermore, indirect content and alternative questions account for 16.7% and 14.3% of their type in the English data. Given this function accounts for only 12.4% of all questions in English, these forms are unusually frequent here. In contrast direct questions are unusually infrequent here, with only 3.8% of direct content and 6.8% of direct polar questions occurring in this function in the English data.

Table 6-12 Percentage direct and indirect questions in framing the discourse per type and form in KIAP-EEFS

	Content	Polar	Declarative	Alternative
%English direct questions	3.8	6.8	0	11.1

%English indirect questions	16.7	36.9	0	14.3
%French direct questions	13	16.7	100	25
%French indirect questions	8.8	6.3	0	0
%Spanish direct questions	28.6	0	0	25
%Spanish indirect questions	5.6	18.2	0	25

The following indirect polar question exemplifies how indirect questions are used to frame the discourse:

One question we are interested in is whether a simple approach to calculating the poverty-efficient allocation of aid yields results similar to a more sophisticated approach. (engecon24)

The example indirect question uses “interested in” and “whether” to launch the question. This question is the main research question of the paper and as such is answered through the delivery of the paper.

For English alternative questions, of which both direct and indirect are worthy of discussion, the following examples exemplify them framing the discourse:

Does DFI ease or exacerbate domestic firms' credit constraints? (engecon41)

Rather, our analysis can be viewed as asking whether fluctuations are primarily caused by "generic" supply or demand shocks and whether those shocks are primarily temporary or permanent in the way they are defined below. (engecon22)

Overall, direct content questions are most characteristic of questions in engecon, as presented in Section 5.5. Therefore, the use of indirect questions in the English data for this functions is evidence of a functional difference that diverges from overall patterns of use.

French questions used for this function do not vary from the overall pattern, presented in Section 5.5, with the tendency in this function to use direct polar questions, such as the following:

Les opportunités d'emploi sont-elles plus grandes pour les plus qualifiés? (frecon06)

In this example, the question refers directly to the research question of the paper. French, like English, also makes use of alternative questions in this function. However, it is worth noting that this question form is so infrequently used that the usage, indicated in Figure

6-3 accounts, for 25% of all direct alternative questions used in French. Such questions pose the research question that guides the paper, giving possible answers in the question:

L'avant-guerre marque-t-elle la fin d'une époque où une phase de transition posant les jalons de la société salariale, aujourd'hui menacée? (frecon11)

Finally, for French question forms, and quite distinct from English and Spanish, there is evidence of declarative questions. These appear like indirect questions which are post tagged with a question mark:

[...] nous nous demandons quel partage des risques en consommation est atteint par les ménages ruraux au Pakistan et quels marchés « d'assurance » leur permettent de s'assurer complètement ou partiellement? (frecon29)

This question is declarative as it does not take an interrogative form and without the question mark, it may have been seen as an indirect question. However, this question also contains content elements with *quel* and alternative elements with *ou*. That being said, fundamentally it was categorised as declarative, as in terms of form, those words were embedded within a declarative sentence. While this example is quite complex, it is rare. In fact, this is the only example of such a question in all the French data. Therefore, it is more of an outlier than a characteristic of this question function.

Finally, with regard to Spanish, once again there is a notable difference. Spanish, though having relatively fewer questions, behaves similarly to English with a spread of direct and indirect questions across, primarily content and polar forms. However, Spanish uses mainly direct content questions for this function. While these are the most frequent questions in Spanish in general, this remains a noteworthy characteristic of questions that frame the discourse, given their unusually frequent occurrence. As mentioned in Table 6-8, this function accounts for 17.3% of all the Spanish questions studied; however, it contains 28.6% of all the direct content questions identified in the Spanish subcorpus in KIAP-EEFS. That being said, this function also accounts for 18.2% of all indirect polar questions in Spanish, meaning it shares this use, albeit to a lesser degree, with English. The following content questions exemplify Spanish's use of this function:

¿Cómo podemos inmunizar contra este tipo de riesgo el valor presente de nuestros flujos esperados? (specon20)

It is also worth noting that the use of alternative questions in this function is similar to English, with both direct and indirect alternative questions being used:

¿En los modelos keynesianos el desempleo es de tipo clásico o de tipo keynesiano?
(specon25)

[...] investigar [...] si estos modelos son complementos o sustitutos en la explicación
de la presencia de crisis. (specon37)

The use of alternative questions in this function allows writers to raise the research question and position the answer proposed by the paper as one of two alternative options. However, again it should be noted that these question forms are very infrequent, relative to other question forms. In fact, 25% of all Spanish direct alternative questions and 25% of all Spanish indirect alternative questions are found in this function. Overall, direct content questions emerge as a particularly frequent characteristic of questions that frame the discourse in Spanish.

6.3.3.1 IFID for indirect questions that frame the discourse

The IFIDs for questions used to frame the discourse demonstrate some interesting differences between each language and the overall patterns identified in KIAP-EEFS. Table 6-13 illustrates how exactly these differences occur in terms of frequency and proportional distribution within KIAP-EEFS.

Table 6-13 IFIDs that generate questions that frame the discourse in WPM and percentage of all questions in KIAP-EEFS

IFID	English	%English	French	%French	Spanish	%Spanish
Noun	10	25	0	0	2	10
Adjective	5	33.3	0	0	0	0
Verb	33	24.1	8	6.4	7	13.6

6.3.3.2 IFID: nouns and adjectives

Nouns in this function account for 25% of noun IFIDs in the English subcorpus and 10% of those in the Spanish subcorpus, rendering them unusually frequent here in English, and relatively infrequent in Spanish. The following examples with the noun ‘question’ illustrate how these indirect questions are posed (for example, where the noun question/*cuestión* followed by whether/*si* raises a polar question):

In this paper we are concerned with the question of whether exceptions to the MFN principle should be made within the dispute phase of international agreements. (noun – question; engecon38)

Primero, intentamos poner a prueba simultáneamente los pronósticos principales de dos grupos de modelos de las crisis (los llamados modelos de primera y segunda generaciones de crisis), a fin de investigar si estos modelos teóricos nos han ayudado a entender por qué ocurren las crisis y elucidar la cuestión de si estos modelos son complementos o sustitutos en la explicación de la presencia de crisis (noun – *cuestión*; specon37)

For adjectives, “interested” as part of the “be interested in” construction reflects a need for discovery; a question the paper intends to answer, as in the subsequent example. The adjectives used for this purpose are infrequent yet account for 33% of all adjectives studied in the English subcorpus.

For testing the productivity of schooling directly, we might be interested in seeing what the data can tell us about the form of (S), and that is one aim of what follows. (adjective – interested; engecon15)

6.3.3.3 IFID: verbs in indirect questions

Verbs are frequently used to raise questions that frame the discourse in English. In fact, 24.1% of all IFIDs that are verbs are found here. This is not apparent for French and Spanish which use mainly direct questions in this function. This renders IFIDs less frequent for these languages in this function. Many of the examples already presented in this section show questions raised through verbal IFIDs, such as:

Finally, we briefly characterize some stylized facts regarding our estimated stocks and ask whether there are trends in net foreign assets and shifts in their composition over time. (verb – ask; engecon02)

Dans l'étude, on a utilisé ces données pour déterminer dans quelles proportions l'analyse longitudinale remettait en question les estimations habituelles. (verb – *déterminer*; frecon01)

Table 6-14 shows the verbs used to set up indirect questions for questions that frame the discourse.

Table 6-14 Verbs that generate questions that frame the discourse in WPM in KIAP-EEFS

Rank	English		French		Spanish	
1	Examine	10	Savoir	5	Investigar	2
2	Explore	8	Montrer	2	Precisar	2

3	Ask	5	Déterminer	2	Explicar	2
4	See	5				
5	Decompose	2				
6	Determine	3				

Owing to English's increased use of indirect questions for this function, there are more examples of recurring verbs being used. English makes most use of “examine”, “explore”, “ask” and “see” illustrated in the following examples.

We also examine how the level of the interchange fee between issuers and acquirers affects consumers' and merchants' behavior, discussing the optimal setting of such an interchange fee from private and social perspective. (engecon29)

We explore whether FPRs play a more important role in protecting knowledge that is transferred outside the source country and firm relative to knowledge inside the source country and firm. (engecon04)

Finally, we briefly characterize some stylized facts regarding our estimated stocks and ask whether there are trends in net foreign assets and shifts in their composition over time. (engecon02)

The aim of the present paper is to see whether the change of currency regime slowed the growth of the volume of Anglo–Irish trade, *ceteris paribus* (engecon28)

In fact, 50% of uses of “examine”, 41% of uses of “explore”, 20% of uses of “ask”, and 50% of uses of “see” to set up indirect questions occur in this function, rendering “examine”, “explore”, and “see” particularly pertinent.

In French however, the only verb worth considering is *savoir*. 33% of all uses of *savoir* occur in this function, as in the following example:

Dans cette optique, la question fondamentale posée aux individus est de savoir combien ils sont prêts à payer pour réduire leur risque l. (frecon08)

This question acts in much the same way as a direct question that frames the discourse, as it refers to the research question guiding the paper. For Spanish, there are no patterns of note here as there are too few indirect questions.

6.3.4 Framing the discourse: location

In terms of the location of questions that frame the discourse and set up research questions, a characteristic that emerges is their frequent use in the introduction section in each language, as Table 6-15 illustrates.

Table 6-15 Framing the discourse question location in WPM and percentage of all questions in KIAP-EEFS

Section	English	%English	French	%French	Spanish	%Spanish
Introduction	28	26.8	33	34.2	39	55.2
Middle	28	8.3	5	2.5	2	1.6
Conclusion	8	16.7	0	0	0	0

As established in Section 5.6, most questions occur in the middle sections of research articles. However, questions in this function challenge this view as questions used to frame the discourse occur mainly in the introduction sections, as reported in Table 6-15. In fact, questions in this function account for 26.8%, 34.2%, and 55.2% of all questions in the introduction sections in English, French, and Spanish respectively. In French and Spanish, questions used to frame the discourse occur minimally in middle sections and do not occur in conclusions. In the case of English, questions are also used to frame the discourse in the middle sections, but this only accounts for 8.3% of all questions in middle sections. For conclusions in English, questions account for 16.7% of all questions in conclusions. This means that there is a salient relationship between this function and introduction and conclusion sections in English. For those questions that occur in the middle section in English, it is worth noting that they also serve to highlight research questions:

Building on these predictions, we examine whether there is a negative relationship between the strength of FPRs and labor flows from the source country to the affiliates. (engecon04)

Building on this prediction, we examine whether knowledge flows to affiliates (relative to other factor flows) are positively related to the strength of FPRs (engecon37)

6.3.5 Framing the discourse: passivity, tense and verbal modality, and sentence type

6.3.5.1 Passivity

For the most part, questions in this function reflect overall patterns of use, *i.e.* they do not use passivity. As presented in Table 6-16, English and French also contain examples of unclear passivity, *i.e.* the IFIDs²³ reflect a non-finite and unvoiced conjugation such as the present participle used in “asking”:

Rather, our analysis can be viewed as asking whether fluctuations are primarily caused by "generic" supply or demand shocks and whether those shocks are primarily temporary or permanent in the way they are defined below (engecon22)

Or the use of *étudier* in the infinitive form here:

L'objectif de cet article est d'étudier les liens entre carrières salariales et droits à la retraite comme dans Caussat [1996] en étendant les microsimulations au secteur public (frecon16)

Both of these are not passive structures, but they are not classifiable as active structures either and are therefore categorised as unclear.

Table 6-16 Passivity in framing the discourse in WPM and percentage of all questions in KIAP-EEFS

Passivity	English	%English	French	%French	Spanish	%Spanish
Not passive	50	12.9	28	14.2	34	19.7
Passive	0	0	0	0	7	21.4
Unclear	13	14.7	10	8.9	0	0

While Section 5.7.1 shows that Spanish rarely shows passivity in questions overall, 21.4% of questions that contain passivity in Spanish occur in this function with 0% in English and French occurring in this function. In this case, Spanish behaves differently with questions like:

¿Qué tipo de relaciones se establecen con la teoría “clásica” típica? (specon25)

²³ Note that, as discussed in Chapter 4, Section 4.3, for direct questions the question verb and for indirect questions the IFID verbs were studied to determine the passivity, and the tense and modality of the question only. As questions can contain multiple clauses there was a possibility of multiples tenses, voices etc. Therefore, the analysis was delimited as such for consistency.

6.3.5.2 Tense and verbal modality

With regard to tense and modality, the present tense is used most frequently in each language, as seen in Table 6-17.

Table 6-17 Tense and verbal modality in framing the discourse in WPM and percentage of all questions in KIAP-EEFS

Tense	English	%English	French	%French	Spanish	%Spanish
Past	3	11.1	0	0	0	0
Present	43	14.4	20	18.2	24	23.8
Future	0	0	0	0	0	0
No tense	13	13.9	10	8.7	10	10.5
Modality	5	6.5	8	11.1	7	23.1

There is evidence of modal verb use in this function with questions using verbs like “might” to show possibility in English:

For testing the productivity of schooling directly, we might be interested in seeing what the data can tell us about the form of (S), and that is one aim of what follows. (frecon15)

There is also evidence of modality surrounding possibility and permission in French with the use *pourraient permettre*:

Dans quelle mesure la réduction de la taille du système par répartition et le développement de la capitalisation pourraient permettre d'augmenter le bien-être de certaines générations sans n'en pénaliser aucune ? (frecon09)

Likewise with Spanish, modality is evoked in the use of *podemos* to denote ability:

¿Cómo podemos inmunizar contra este tipo de riesgo el valor presente de nuestros flujos esperados? (specon20)

For English, this function only captures 6.5% of all modality identified in this study. French, slightly higher proportionally, captures 11.1% of all modals used in posing French questions. Most significantly, 23.1% of all questions posed with modality in Spanish occur within this function, rendering it relatively frequent. However, it is important to keep in mind that while they may be interesting examples, there are not sufficient examples to consider them a characteristic feature of questions that frame the discourse.

6.3.5.3 Sentence type

Finally, with regard to sentence type, there are some differences to be seen as indicated in Table 6-18 and Figure 6-5.

Table 6-18 Percentage sentence type in framing the discourse in KIAP-EEFS

Sentence type	English	French	Spanish
%Simple	1.5	17.2	25
%Complex	20	8.3	0
%Compound	8.7	33.3	30
%Compound-complex	25.9	14.3	20

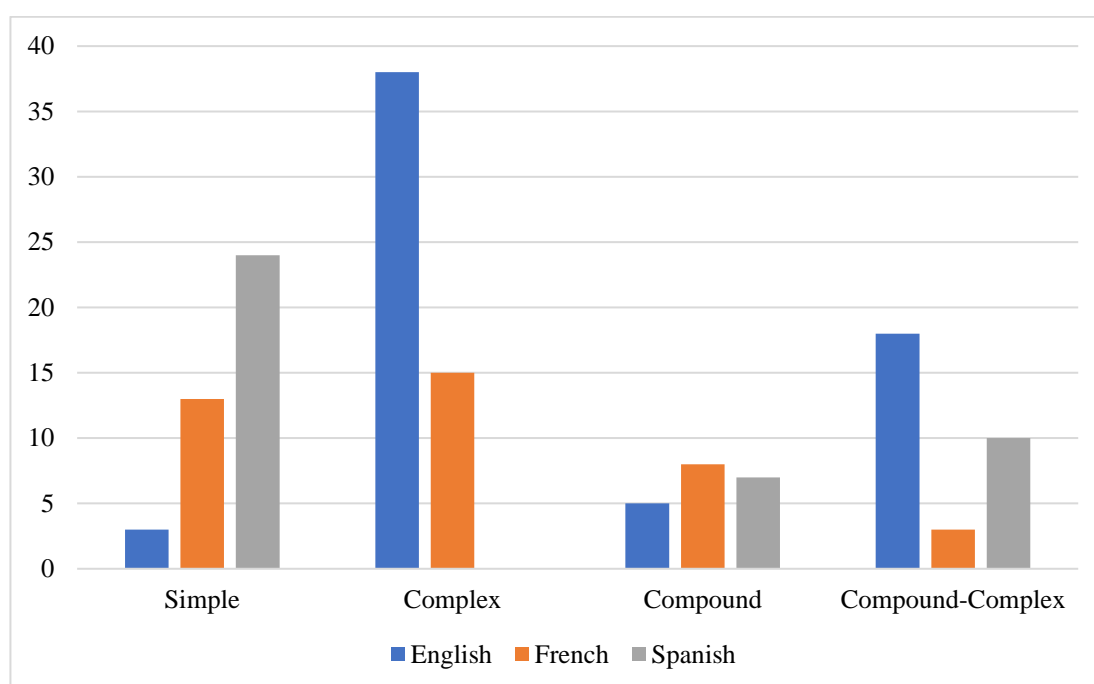


Figure 6-5 Sentence types in framing the discourse in KIAP-EEFS in WPM

First, while Section 5.7.3 shows that English uses many simple sentences, this function makes greater use of complex sentences in English. In fact, only 1.5% of the simple sentence questions in English occur in this function. This is remarkably low for a function that accounts for 12.3% of all the questions studied in English. Instead, English uses more complex and compound-complex sentences which account for 20% and 25.9% of such sentence types identified.

French sentence types in questions that frame the discourse seem to mirror the overall patterns presented in Section 5.7.3. However, upon further inspection, it can be seen that simple and compound sentences are unusually frequent here. This function accounts for 11.6% of all questions studied yet it contains 17.2% of all simple sentence questions and 33.3% of all compound sentence questions identified in French, and only 8.3% of the complex sentences identified.

Finally, Spanish uses more simple sentences than any other language for this function; they account for 25% of all simple sentence questions in the Spanish data, rendering simple sentences unusually frequent for this function. Therefore, as in the case of French, simple sentences are a notable characteristic of Spanish questions that frame the discourse. Moreover, the lack of complex questions is also noteworthy. This is because this section contains 0% of Spanish complex questions despite accounting for 17.3% of all Spanish questions and despite the fact that 24.5% of all questions in the Spanish data are complex. Otherwise, this function captures large proportion of all compound and compound-complex sentence questions, as 20% of both of these question sentence types occur in this function.

6.3.6 Framing the discourse: conclusion

For framing the discourse, the following findings outline the most frequent characteristics of questions:

1. English and French questions that frame the discourse are long while Spanish questions are short.
2. English uses mainly indirect questions and French and Spanish uses direct questions.
3. English uses mainly indirect polar questions, French uses mainly direct polar questions and Spanish uses mainly direct content questions.
4. These questions occur in the introduction and middle sections in English and in the introduction sections in French and Spanish.
5. There is little evidence of passivity in any language for this function.
6. The present tense is most common in all languages.
7. English and French use mainly complex sentences, and Spanish uses mainly simple sentences.

When considering questions that frame the discourse, some distinctive characteristics emerge in each language when compared to other functions:

1. Questions are unusually long in English and French while Spanish question lengths reflects overall patterns in question length in Spanish as presented in Chapter 5.
1. Indirect questions are unusually frequent in English and direct questions are unusually frequent in French and Spanish.
2. English uses an unusually frequent number of indirect polar questions, indirect content questions, and indirect alternative questions here. French uses an unusually frequent number of direct polar and direct content questions, and Spanish uses an unusually frequent number of direct content and indirect polar questions to frame the discourse.
3. These questions are unusually frequent in introduction sections in each language.
4. Spanish exhibits and unusually frequent use of passivity for this function while English and French do not.
5. English and French exhibit an unusually frequent use of present tense while Spanish makes unusually frequent use of both present tense and verbal modality.
6. English makes salient use of complex and compound-complex sentences while French makes use of simple and compound sentence types for questions and Spanish use more compound, simple, and complex-compound questions in decreasing order.

6.4 Organising the text

6.4.1 Organising the text: frequency

In terms of frequency, questions that organise the text account for around 40% of all questions in each language, as seen in Table 6-19. This makes this function the most frequent function in each language.

Table 6-19 Questions that organise the text in WPM and percentage of all questions in KIAP-EEFS

	English	French	Spanish
Total in WPM	199	130	97
Percentage %	39	40	41

6.4.2 Organising the text: sentence length

Each language appears to follow overall length of questions presented in Section 5.4, with English and Spanish having shorter questions than French. While the average length of questions in Spanish is somewhat longer in Table 6-20, this is due to Spanish having some outliers in the higher question lengths with questions containing 50+, 60+, and 80+ words:

Nos interesa precisar si las exportaciones de estos países las justifican factores tradicionales basado en el aprovechamiento de la riqueza natural de cada país, con predominio de los productos agrícolas y minerales, cuyas expectativas futuras parecen escasas de acuerdo con el comportamiento mostrado por la demanda en dichos mercado, o si, por el contrario, son resultado de un proceso de industrialización en marcha. (specon02)

Table 6-20 Organising the text question length statistics in KIAP-EEFS

Sentence length	English	French	Spanish
Average question length in words	18.9	25.3	20.9
Median question length in words	19	21	17
Mode question length in words	19	16	7

Table 6-21 Organising the text question length in WPM and percentage of all questions in KIAP-EEFS

Sentence length	English	%English	French	%French	Spanish	%Spanish
15 words or fewer	85	34.3	28	33.3	46	36.6
16-25 words	68	40	58	43.4	29	52.2
26 words or more	45	38.3	45	41.9	22	39.1

When considering the relative frequency of sentence lengths per function, this function makes unusually infrequent use of shorter questions in each language where only 34.3%, 33.3%, and 36.6% of all shorter questions occur in this function in English, French and Spanish. This renders this length less salient here. This is because this function accounts for 39%, 40%, and 41% of questions in English, French, and Spanish which means that less than 39%, 40%, and 41% of short questions in English, French, and Spanish occur here. Therefore, the fact that there are many short questions in this function appears to be

due to the fact that there are many short questions overall. Proportionally, however there is nothing noteworthy about the relationship between short questions and this function, other than them being slightly less frequent than may be expected. For English 40% of questions of middling length occur here. Furthermore, 43.4% and 52.2% of questions of middling length in French and Spanish, respectively, occur in this function. Therefore, when considering the proportional distribution of questions within and across each language, a clear feature of this function is the shared use of questions of 16-25 words.

6.4.3 Organising the text: question type and form

English and Spanish use both direct and indirect questions, while French uses mainly indirect questions, as presented in Table 6-22 and Figure 6-6.

Table 6-22 Percentage direct and indirect questions in organising the text in KIAP-EEFS

Question category	English	French	Spanish
%Direct questions	39.2	36	44.6
%Indirect questions	39.5	46.3	33.3

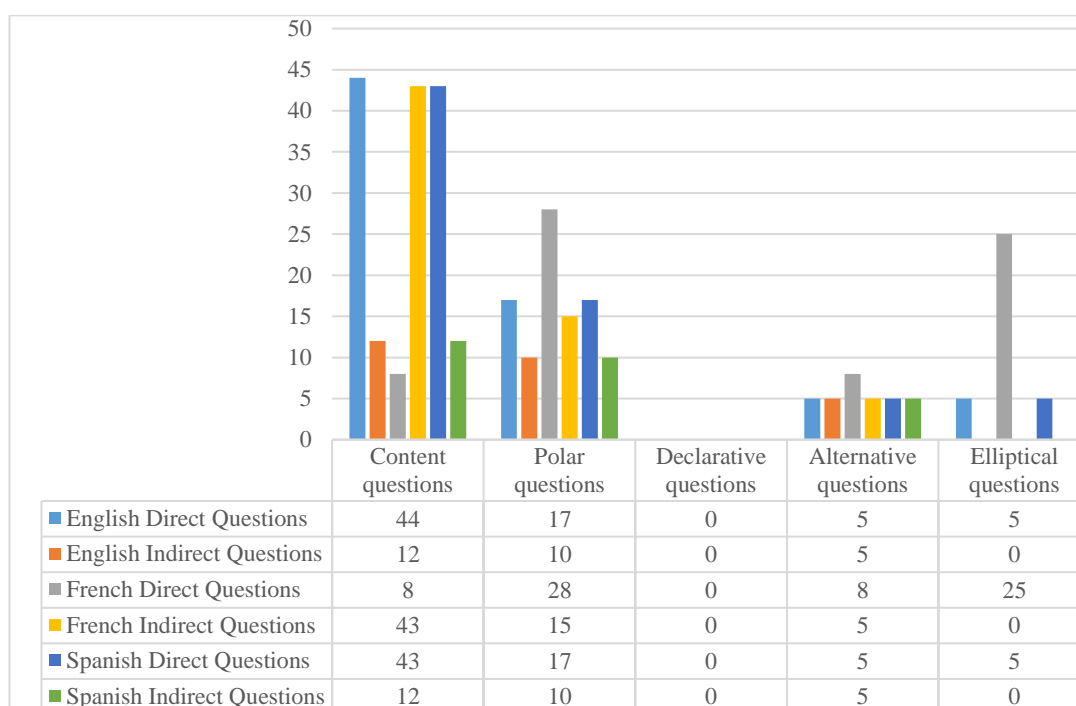


Figure 6-6 Question types and forms in organising the text in WPM in KIAP-EEFS

Relatively, this function captures 39.2% of all direct questions and 39.5% of all indirect questions in English. Given that this function accounts for 39% of questions in English, it is clear that this function does not predict the notable use of one or the other. Rather, it makes proportionate use of direct and indirect questions, and therefore mirrors the overall patterns, described in Section 5.5. French makes more salient use of indirect questions. Only 36% of all direct questions in French occur here, while 46.3% of all indirect questions in French occur in this function. Finally, for Spanish 44.6% of direct questions occur in this function, while only 33.3% of all indirect questions occur here, rendering direct questions in Spanish of note here.

Table 6-23 Percentage direct and indirect questions in organise the text per type and form in KIAP-EEFS

	Content	Polar	Alternative	Elliptical
%English direct questions	31.3	62.1	55.6	20
%English indirect questions	33.3	45.5	42.9	0
%French direct questions	13	30.5	75	90.9
%French indirect questions	50	37.5	50	0
%Spanish direct questions	42.9	58.3	50	28.6
%Spanish indirect questions	27.8	36.4	50	0

As presented in Table 6-23, while content questions do feature in English, direct and indirect content questions are unusually infrequent as this function. This is because this function only accounts for 31.3% and 33.3% of all direct and indirect content questions in English. Polar questions however, tell another story with 62.1% of all direct polar questions and 45.5% of all indirect polar questions found in English being used to organise the text by signalling what is coming next:

Do strong FPRs increase the location of knowledge factors outside the source country?
(engecon4)

Finally, we examine whether the skill premium tended to rise in country-decades with SBTC concentrated in skill-intensive sectors and fall in country-decades with SBTC concentrated in unskill-intensive sectors. (engecon23)

Therefore, polar questions have a particularly strong relationship with this function in English. Only 20% of all direct elliptical questions in the English data occur in this function, rendering them quite infrequent. Alternative questions however, appear

unusually frequently with 55.6% of all direct and 42.9% of all indirect alternative questions in English occurring within this function. These questions can appear at the end of a section to indicate the contribution of the following section:

Is the net effect of FPRs on US ownership advantage a market expanding effect or a market power effect? (engecon4)

French like English makes frequent use of both direct and indirect questions. 46.2% of all indirect questions in French occur in this function, as opposed to only 36% of all direct questions. This indicates that indirect questions are unusually frequent in this function in the French data. In terms of direct questions, only 13% of all direct content questions and 30.6% of all direct polar questions occur in this function, rendering them unusually infrequent. Conversely, 50% of all indirect content questions occur in this section to organise the text, as in the following question:

La question qui se pose alors est de savoir quel sera le " prix " de cette action évalué par les différents individus. (fecon08)

Through the use of the indirect question with *alors*, for example, the reader is signalled as to what will be discussed next. These questions are unusually frequent and emerge as a characteristic of this function. Polar questions also occur frequently in this function in French; however relative to their overall use, this function only includes 30.6% of all direct polar questions and 37.5% of all indirect polar questions, rendering them slightly less frequent for this function. Another important distinction for French is the use of elliptical questions, which although not the most frequent, account for 90.9% of all elliptical questions in French. This means that there appears to be a stronger relationship between this function and elliptical questions than direct content questions, for example. Such elliptical questions, as in the following examples, occur largely within subtitles and indicate the questions to which the subsequent section will attend:

1.2. Le consommateur, producteur de soutenabilité? (frecon02)

It is also worth noting that, although very infrequent, there is evidence of alternative questions occurring in this function, which account for 75% of all direct alternative and 50% of all indirect alternative questions in French. These questions organise the text by signalling to the reader the two or more possible answers to questions, which are usually delivered in the sections that follow the question, for example:

La rémunération du directeur doit-elle être considérée comme un salaire ou comme une part du profit? (frecon18)

Si ces résultats devaient toutefois être confirmés, il serait légitime de se demander s'ils reflètent la spécificité du dispositif français de formation continue ou si leur portée est plus Générale. (frecon21)

Finally, for Spanish questions that organise the text, indirect questions appear uncharacteristic of this function with only 27.8% of all indirect content questions and 36.4% indirect polar questions being used here. These are unusually infrequent, given this function accounts for 41% of all Spanish questions. What emerges as more characteristic of this function is the use of direct content questions which account for 42.9% of all direct content questions in specon. These are used like English and French to signal to the reader the topics which will be addressed, as demonstrated by the following example of a question in subtitle:

1. ¿Qué es la globalización? (specon01)

Direct polar questions are even more noteworthy in Spanish. In fact, within this function alone, 58.3% of all direct polar question in Spanish occur. These questions are used to organise the text in subtitles and signal what the subsequent section will address, the questions it will answer, as in the following example:

¿Quiere esto decir que las necesidades están jerarquizadas? (specon22)

In terms of lesser used questions, Spanish, like English, makes use of some elliptical questions. Unlike French however, this function only accounts for 28.6% of all elliptical questions in Spanish. Therefore, there is not enough evidence to indicate that this is a particular trait of questions that organise the text. However, this function does account for 50% of all direct and indirect alternative questions in Spanish, and while they are not a particularly frequent feature of Spanish questions overall, they are used to organise the text by pointing the reader to possible outcomes, as the following example indicates:

¿La demanda de dinero en Keynes es una demanda en términos nominales o reales? (specon25)

Overall, for this function it is clear that each language differs slightly from the others and from the overall patterns presented in Section 5.5.

6.4.3.1 IFID for indirect questions that organise the text

English, French and Spanish use different IFIDs to raise questions that organise the text in KIAP-EEFS, as indicated in Table 6-24. Verbs are most frequent in this function, followed by nouns, and the sole use of adjectives in English.

Table 6-24 IFIDs that generate questions that organise the text in WPM and percentage of all questions in KIAP-EEFS

IFID	English	%English	French	%French	Spanish	%Spanish
Noun	18	43.8	5	33.3	7	30
Adjective	8	50	0	0	0	0
Verb	50	37	58	48.9	19	36.4

6.4.3.2 IFID: nouns and adjectives

In the English data, both nouns and adjectives feature relatively highly, with this function accounting for 43.8% of all nouns and 50% of all adjectives in English. However, it should be noted that adjectives are infrequent overall. The noun “question” recurs several times in English to pose a question, as well as the adjective clear:

Although violation of the condition presented in Proposition 2 implies that we cannot be in Regime A, it leaves open the question of which set of information constraints may bind at the optimum. (engecon32)

But to separate trade from technology we need to be clear how technology affects trading opportunities and thus the factor contents. (engecon10)

Both examples show questions that will reveal some information. This signals to the reader the focus of the upcoming discussion and therefore organises the text.

For French and Spanish, this function makes slightly less use of nouns with only 33.3% of nouns occurring in French and 30% in Spanish. That being said, given that nouns are so infrequent in French and Spanish, there is hardly sufficient data to make any claims of substance. Those nouns that appear are *question/cuestión/pregunta* which reflect overall patterns of use, described in Section 5.5.2.

6.4.3.3 IFID: verbs in indirect questions

Looking more closely at verbs, the verbs most frequently used in each language to set up questions are presented in Table 6-25. In English the verb “tell” is most common in this function, accounting for 100% of the examples of “tell” in the English data. For organising the text, “examine” and “ask” account for 37.5% and 30% of all examples of these verbs used to set up questions in English. This makes the relationship between “tell” and this function more salient.

Table 6-25 Verbs that generate questions that organise the text in WPM in KIAP-EEFS

Rank	English		French		Spanish	
1	Examine	8	Savoir	10	Preguntar	5
2	Ask	8	Déterminer	10	Evaluar	5
3	Tell	8	Montrer	8		
4	Explore	5	Interroger	8		
5	Check	5	Tester	8		
6			Demander	5		

The following example shows how “tell” is used to organise the text by moving the analysis forward. This is done by asking questions and discussing how to answer them:

However, it does not tell us what the levels of the transfers should be. (engecon32)

For French, *savoir* and *déterminer* are both examples of question words that feature more prominently in this function than others. 66% of all uses of these verbs in are used to organise the text. *Tester* is also of note here as it only occurs in this function. These question words are effective at creating questions that signal to the reader what to expect next in the text, such as:

L'objectif sera précisément de savoir s'il existe une politique permettant de supprimer ce système par répartition en un nombre fini de périodes. (frecon09)

Il semble donc intéressant de pouvoir tester si les ménages participant à des contrats de métayage parviennent à mieux s'assurer contre les risques que les autres. (frecon29)

Notre objectif est maintenant de déterminer les prix d'équilibre. (frecon43)

Finally, for Spanish, there is little variety here with the use of *preguntar* which accounts for 40% of all examples of *preguntar* that set up indirect questions. This reflects the overall patterns presented in Section 5.5.4. However, the verb *evaluar* only occurs in this function to set up questions that organise the text and acts as a means to signal what the next section will address and how it will do so:

En el presente estudio, el cálculo de las elasticidades está desagregado con el fin de evaluar cómo el diferencial en la reducción de tarifas para las 49 ramas del sector manufacturero ha afectado la estructura del comercio internacional y el crecimiento en cada rama industrial. (specon30)

6.4.4 Organising the text: location

This function is largely found in the thematic middle sections of research papers and it was found that 55.6%, 59.5%, and 60.3% of all questions in English, French, and Spanish questions that occur middle sections, occur in this function. Therefore, this function uses questions in the middle section in an unusually frequent manner while it uses questions in all other sections much less frequently. The distribution across language is shown in Table 6-26.

Table 6-26 Organising the text question location in WPM and percentage of all questions in KIAP-EEFS

Section	English	%English	French	%French	Spanish	%Spanish
Introduction	13	12.2	13	13.3	5	6.9
Middle	186	55.6	118	59.5	92	69.3

6.4.5 Organising the text: passivity, tense and verbal modality, and sentence type

6.4.5.1 Passivity

While there is evidence of some passivity in each language, there is no evidence that this function correlates more strongly with the use of passivity than other functions, as seen in Table 6-27.

Table 6-27 Passivity in organising the text in WPM and percentage of all questions in KIAP-EEFS

Passivity	English	%English	French	%French	Spanish	%Spanish
-----------	---------	----------	--------	---------	---------	----------

Not passive	163	41.9	55	28.6	78	45.1
Passive	10	33.3	5	28.6	12	35.7
Unclear	25	29.4	10	62.2	7	23.1

This is because only 33.3% of passive questions in English, 28.6% of passive questions in French, and 35.7% of passive questions in Spanish occur in this function. For French, it is also worth noting that owing to the frequency of elliptical questions there are several questions categorised as unclear. This has also impacted tense and sentence type for French which is presented in the following two sections.

6.4.5.2 Tense and verbal modality

In the English, French, and Spanish data, the past tense appears to be most frequently used, while in French, there is also evidence of several tenseless questions, as reported in Table 6-28 and Figure 6-7. Despite this frequent use of present tense, there are interesting patterns in the use of past and future tenses, and modal verbs in this function. In fact, 44.4% of all past, 41.5% of all present, and 42.9% of all future questions in English occur in this function. Spanish makes most use of present tense which and this function accounts for 47.6% of all questions in the present tense in specon. Furthermore, although the presence of past and future tenses are very infrequent in Spanish in general, as mentioned in Section 5.7.2, it is worth noting that 100% of past and 50% future questions in specon occur in this function. For French, the use of future tenses is salient, as 100% of questions in the future occur in this function. However, future tenses in French are very infrequent overall. Therefore to claim a relationship between this function and future tenses in French would have be a tentative claim. French also makes greater use of tenseless questions as 63.3% of questions without tense in French occur in this function. This is largely due to the presence of elliptical questions.

Table 6-28 Percentage tense and verbal modality in organising the text in KIAP-EEFS

Tense	English	French	Spanish
Past	44.4	33.3	100
Present	41.5	30.6	47.6
Future	42.9	100	50

No tense	33.3	65.2	34.2
Modality	35.5	14.8	23.1

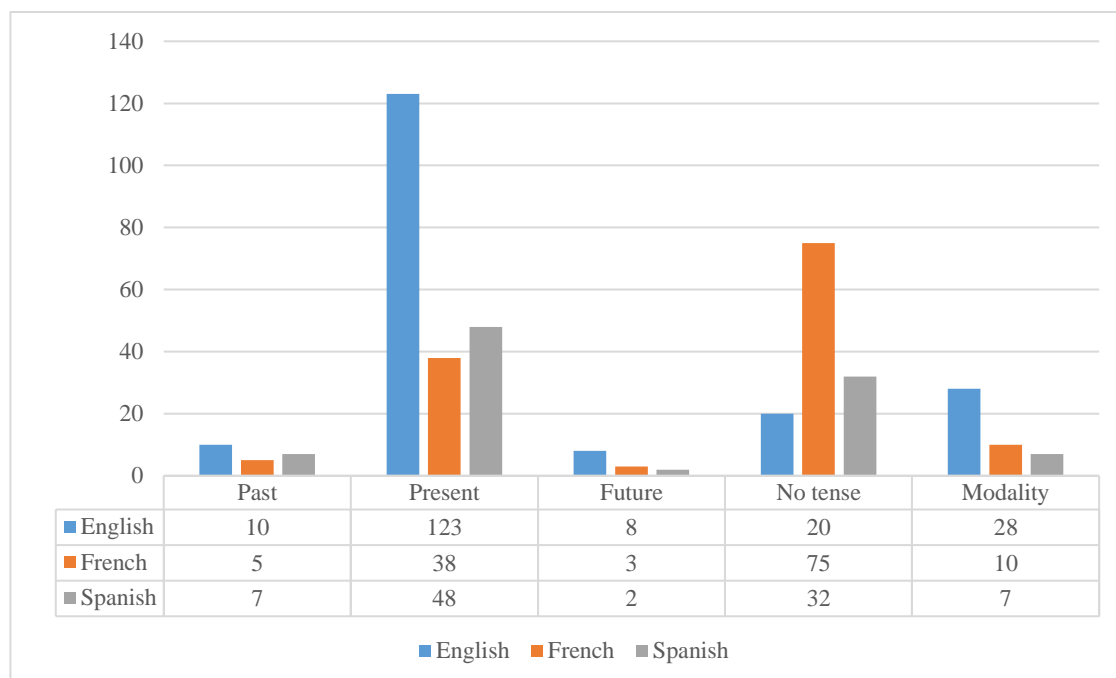


Figure 6-7 Tense and verbal modality in organising the text in KIAP-EEFS in WPM

6.4.5.3 Sentence type

Finally, for sentence types Table 6-29 and Figure 6-8 show that English and Spanish questions use mainly simple and complex sentences and French questions use mainly complex and no clause sentences. However, in English, neither are particularly significant. This because 37.3% of all simple and 40% of all complex questions occur in this function, which is a finding that is consistent with the overall patterns of use, described in Section 5.7.3. Instead, of note for English is the use of compound sentence questions. Despite compound sentence questions not being very frequent in the data overall, as seen in Section 5.7.3, they are salient in this function as 52.2% of all compound sentence questions serve to organise the text in English. For Spanish simple sentences also follow overall trends with 40% occurring in this function. However, complex sentences do appear to be quite salient, with 58.3% of all complex questions occurring in this function in Spanish.

Table 6-29 Percentage sentence type in organising the text in KIAP-EEFS

Sentence type	English	French	Spanish
Simple	37.3	13.8	40
Complex	40	43.1	58.3
Compound	52.2	33.3	30
Compound-complex	29.6	42.9	30
No clause	37.5	91.7	25

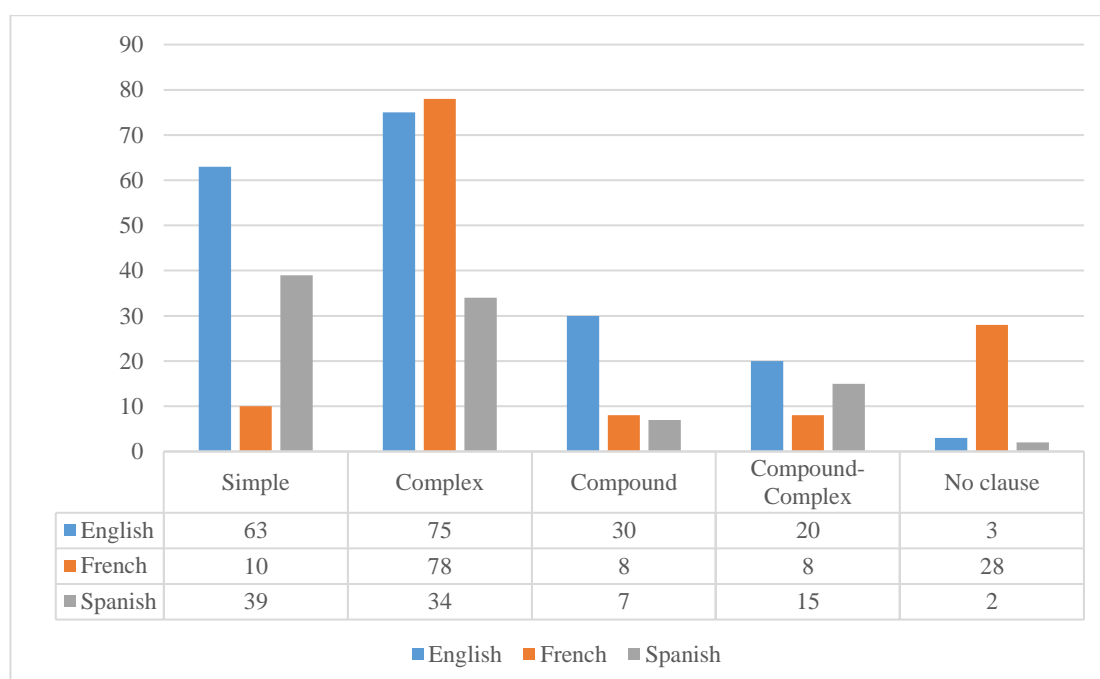


Figure 6-8 Sentence types for questions that organise the text in KIAP-EEFS in WPM

For French, complex and compound questions are particularly salient for organising the text. This is evidenced in the finding that 43.1% of all complex and 42.9% of all compound questions occur in this function in French. Finally, French also makes frequent use of no clause sentences, due largely to elliptical questions. As such, this function contains 91.7% of all clauseless questions in French.

6.4.6 Organising the text: conclusion

For organising the text, the following findings are noteworthy in terms of most frequent characteristics of questions studied:

1. English and Spanish uses shorter questions and French questions are middling in length.
2. English and Spanish use mainly direct questions while French uses a combination of direct and indirect questions.
3. English and Spanish use mainly direct content questions and French uses mainly indirect content questions.
4. All three language see these questions occur in the middle sections.
5. There is no evidence of passivity as being frequent in this function.
6. Each language uses mainly the past tense to pose questions.
7. English and French use mainly complex questions, while Spanish uses simple questions.

When considering questions that organise the text, some distinctive characteristics emerge in each language when compared to other functions:

1. Short questions are unusually infrequent in this function in each language and questions of middling length are unusually frequent in French and Spanish here.
2. English follows the overall patterns of use of direct and indirect questions while indirect questions are unusually frequent in French and direct questions are unusually frequent in Spanish.
3. English makes unusually frequent use of direct and indirect polar questions, as well as indirect alternative questions. French makes unusually frequent use of indirect content questions and direct elliptical questions, and Spanish makes unusually frequent use of direct polar and direct and indirect alternative questions.
4. These questions are unusually frequent in middle sections
5. There is no salient use of passivity
6. English makes unusually frequent use of past tenses, French makes unusually frequent use of future tense, and Spanish makes unusually frequent use of past, future and present tenses in decreasing order.
7. English makes salient use of compound sentence types, French makes salient use of complex and complex-compound sentence types, and Spanish makes salient use of complex sentence types for these questions.

6.5 Creating a niche

6.5.1 Creating a niche: frequency

Questions that create a niche account for 10.9% of all questions identified in French, rendering them somewhat more pertinent to the French data. While they remain relevant

for the English and Spanish data, they do account for a slightly smaller share of questions at 8.5% and 8.2% respectively, as seen in Table 6-30.

Table 6-30 Questions that create a niche in WPM and percentage of all questions in KIAP-EEFS

	English	French	Spanish
Total in WPM	43	35	19
Percentage %	8.5	10.9	8.2

6.5.2 Creating a niche: sentence length

English and French questions that create a niche mirror overall question length, described in Section 5.4. That is to say, the English questions are shorter with 11.1% of questions of 15 words or fewer occurring in this function. Conversely, French questions are largely in the mid-length category with 13.2% of questions of 16-25 words occurring in this function, as can be seen by Tables 6-31 and 6-32. However, Spanish questions that create a niche are much longer than most Spanish questions in the data, as 21.7% of all questions of 26 words or more occur in this function. This makes these specified lengths unusually frequent in each language and makes longer questions a characteristic of Spanish questions that create a niche.

Table 6-31 Creating a niche question length statistics in KIAP-EEFS

Sentence length	English	French	Spanish
Average question length in words	16.3	21.8	30
Median question length in words	14	20.5	30.5
Mode question length in words	5	16	N/A

Table 6-32 Creating a niche question length in WPM and percentage of all questions in KIAP-EEFS

Sentence length	English	%English	French	%French	Spanish	%Spanish
15 words or fewer	28	11.1	8	9.1	7	5.7
16-25 words	3	1.8	18	13.2	0	0

26 words or more	13	10.6	10	9.3	12	21.7
------------------	----	------	----	-----	----	------

6.5.3 Creating a niche: question type and form

Direct questions occur frequently in English and French, and somewhat less frequently in Spanish. Indirect questions are most common in French and Spanish in this function. In English, direct questions are a strong feature for this function, as shown in Table 6-33 and Fig 6-9. While this function accounts for 8.5% of all questions in English, it accounts for 10.4% of direct questions but only 5.3% of indirect questions. Therefore, in English, indirect questions appear unusually infrequent. For French, direct questions feature in this function, although they are not noteworthy. More interestingly, 13% of all indirect questions in French occur in this function, rendering them unusually frequent. For Spanish, direct questions are unusually infrequent, as only 4.6% of all direct questions occur in this function. In parallel, 15.2% of all indirect questions occur in this function in Spanish.

Table 6-33 Percentage direct and indirect questions in creating a niche in KIAP-EEFS

Question category	English	French	Spanish
%Direct questions	10.4	9.3	4.6
%Indirect questions	5.3	13	15.2

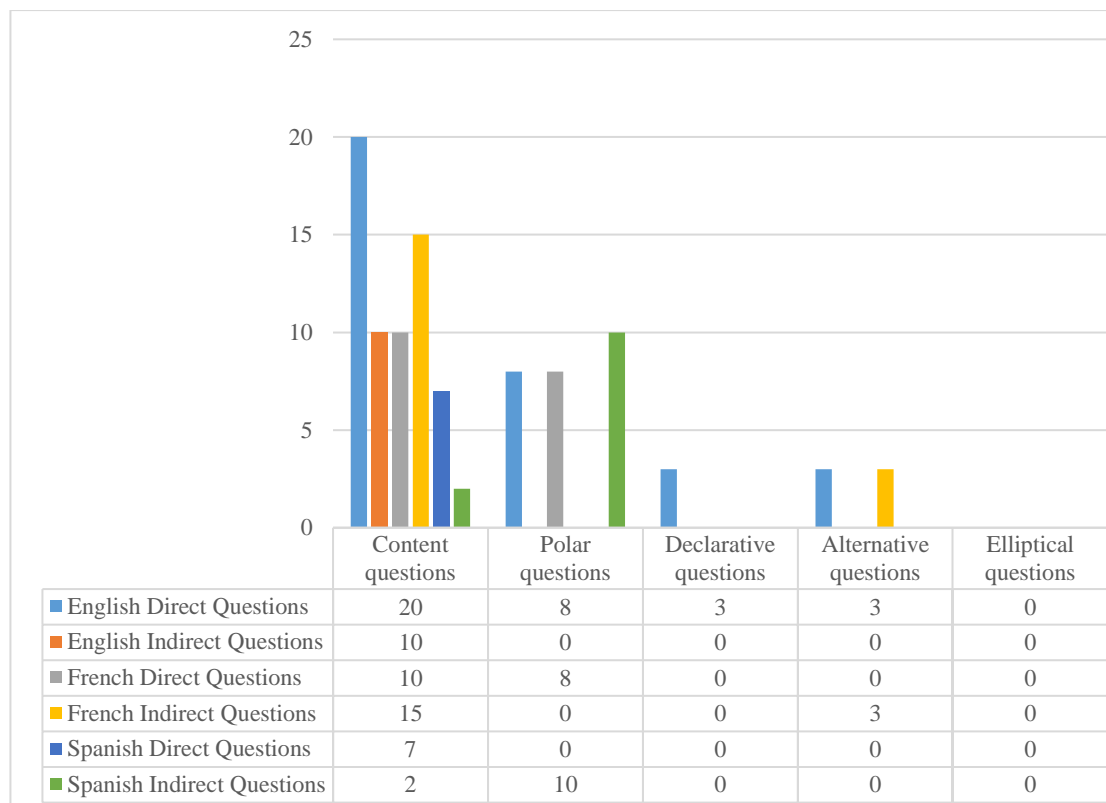


Figure 6-9 Question types and forms in creating a niche in WPM in KIAP-EEFS

In English, direct content and polar questions occur frequently in this function. When compared to their presence throughout the English data, this functions accounts 10% and 10.3% of direct content and polar questions, as seen in Table 6-34. These questions highlight a niche, for example:

How large are these gains? (engecon06)

In this example, this content question uses a question to claim centrality as following this question, the author identified something lacking in the literature. Similarly, in the following example, a polar question is used in the paper to point to a gap in the literature that the research article aims to address and attenuate:

Should they be harmonized across countries? (engecon25)

English also also uses declarative and alternative questions for this function. Although they appear to be somewhat infrequent, this function includes 50% of the direct declarative questions and 11.1% of the direct alternative questions identified in English.

The following declarative question, for example, raises a question around an established problem that the author claims they can address:

Our model also provides an answer to the questions, why negotiations along the lines of [Coase 1960], in which consumers and merchants negotiate prices that internalize any externalities that exist between them, will not take place when surcharging is allowed? (engecon29)

However, declarative questions are so infrequent in the English data it is difficult to treat this finding as a anything of substance, as there are too few examples to render it characteristic of this function. The same can be said for alternative questions, such as the following example which points to a gap in the literature that the author addresses:

Should those taxes be levied by the destination or origin principle? (engecon25)

Table 6-34 Percentage direct and indirect questions in creating a niche per type and form in KIAP-EEFS

	Content	Polar	Declarative	Alternative
%English direct questions	10	10.3	50	11.1
%English indirect questions	11.1	0	0	0
%French direct questions	17.4	8.3	0	0
%French indirect questions	17.6	0	0	25
%Spanish direct questions	7.1	0	0	0
%Spanish indirect questions	5.6	36.4	0	0

Unlike English, French shows greater distributional variety in the forms of questions. In this function, 17.4% of direct content and only 8.3% of the direct polar questions identified in the French data occur. These questions claim centrality by posing questions that have not been answered elsewhere. The following direct content question, which helps the writer to identify a niche also helps the writer occupy it:

Pourquoi les villes se forment où elles se forment ? (frecon36)

For indirect questions in French, which are relatively more frequent in this function than others, 17.6% of indirect content questions and 25% of indirect alternative questions occur in this function. Given that this function accounts of only 10.9% of all questions in French, these occurrences are unusually frequent. These questions similarly pose questions to identify niches:

L’objet de cet article est d’étudier dans quelle mesure l’opérateur dominant peut avoir une incitation privée à étendre son pouvoir de monopole sur l’infrastructure essentielle vers le segment concurrentiel. (frecon46)

Cependant, on peut se demander s’il faut ou non réguler les prix finaux, alors que ceux-ci sont offerts dans un contexte concurrentiel. (frecon46)

Both of these questions work together to claim centrality by identifying the aim of the article and by problematising the way in which it the goal of the article is achieved. The questions identify issues in the literature and proposes a means for the writer to address them.

Finally, for Spanish this function behaves somewhat more like French than English. Spanish uses mainly indirect polar questions for this function, and 36.4% of all indirect polar questions occur in this function, rendering it unusually frequent as a question form. The following example shows how an indirect question can claim centrality by emphasising the important offering the research article makes.

Uno de los aspectos más importantes de este análisis consiste en saber si la intervención política podría eliminar el trabajo infantil mediante leyes que prohíban o impongan la educación obligatoria (specon39)

6.5.3.1 IFID for indirect questions that create a niche

There is little of interest in this function regarding nouns, with English and Spanish using a very few nouns and French using none at all. For both English and Spanish the noun question/*pregunta* is used, although there is no major pattern here to touch upon. As Table 6-35 shows, verbs are of more interest here.

Table 6-35 IFIDs that generate questions that create a niche in WPM and percentage of all questions in KIAP-EEFS

IFID	English	%English	French	%French	Spanish	%Spanish
Noun	3	6.25	0	0	2	10
Verb	7	5.6	18	14.9	10	18.2

The variety of verbs used in this function when creating an indirect question is illustrated in Table 6-36. There is little of interest here in terms of similarity and difference. Once again, the verb “be” is present in each language, while the remaining verbs are all different

across and within languages. In French, *montrer* is used to indicate what the research will show, hence indicating their centrality and contribution:

Nous montrons comment estimer les paramètres de préférences des ménages sous l'hypothèse nulle par une méthode de variables instrumentales (frecon29)

All instances of *montrer* used to claim centrality, it should be noted, come from the same paper. Therefore, it is not sufficient to say that this usage is characteristic of this function in French. Likewise, in Spanish, the verb *preguntarse* is used in this function to indicate a gap to be addressed:

Además, al preguntarse cuál sería el sistema de tipo de cambio apropiado. (specon16)

However, once again, there is not sufficient evidence to render this a pattern for this function. Within English, there are fewer re-uses of verbs in this function, and no clear patterns emerges as to the relationship between verb and this function.

Table 6-36 Verbs that generate questions that create a niche in WPM in KIAP-EEFS

Rank	English		French		Spanish	
1	Examine	3	Montrer	5	preguntar	5
2	Understand	3	Interroger	3	investigar	2
3			Demander	3	saber	2
4			Étudier	3	conocer	2
5			Comprendre	3		
6			Effectuer	3		

6.5.4 Creating a niche: location

This question function features prominently in introduction sections in each language, as can be seen in Table 6-37. In fact, this function includes 32.5%, 27% and 24.1% of all questions in introduction sections in English, French, and Spanish respectively. These are unusually high frequencies for this function rendering the location a shared characteristic across each language. Moreover, it is worth noting that this question function is unusually

infrequent in middle sections which account for 0.8% of all English, 3.8% of all French, and 1.6% of all Spanish questions in the middle section.

Table 6-37 Creating a niche question location in WPM and percentage of all questions in KIAP-EEFS

Section	English	%English	French	%French	Spanish	%Spanish
Introduction	33	31.7	25	26.3	17	24.1
Middle	3	0.8	8	3.8	2	1.6
Conclusion	7	16.7	3	12.5	0	0

6.5.5 Creating a niche: passivity, tense, polarity and sentence type

6.5.5.1 Passivity and tense and verbal modality

Regarding passivity, there is little evidence of passivity in this function, as can be seen in Table 6-38. Moreover for tenses, present tense and verbal modality are frequent in English, while in French, present tense is most frequent, as presented in 6-39. For Spanish, present tense and tenseless questions occur most prominently. Beyond these more frequent items, there are some interesting findings worth discussing. Given that this function only accounts for 8.5% of questions in the English data, it is noteworthy that it contains 25% of all passive questions, 14.3% of questions using the future tense, and 16.1% of questions using modals:

Should they be harmonized across countries? (engecon25)

Why negotiations along the lines of [Coase 1960], in which consumers and merchants negotiate prices that internalize any externalities that exist between them, will not take place when surcharging is allowed? (engecon29)

While of course, it is clear that passivity, future tenses and modal verbs are rarely used, they are unusually frequent in English in this function. For French, there are no examples of passives despite this function capturing 10.9% of all questions in French. Moreover, there is an unusually frequent use of present tense in this function with 18.4% of all questions in the present being used to create a niche in French, as in the following example:

Nous montrons comment estimer les paramètres de préférences des ménages sous l'hypothèse nulle par une méthode de variables instrumentales (frecon29)

Spanish tells a similar story, with no noteworthy use of passivity. It reflects overall patterns in Sections 5.7.1 and 5.7.2, with limited passivity and the primary use of the present tense. The tenseless items in Spanish in this function derive from the IFID verbs that form non-finite clauses in raising indirect questions.

Table 6-38 Passivity in creating a niche in WPM and percentage of all questions in KIAP-EEFS

Passivity	English	%English	French	%French	Spanish	%Spanish
Not passive	23	5.8	23	11.7	15	8.5
Passive	8	25	5	0	2	7.1
Unclear	13	14.7	8	11.1	2	7.7

Table 6-39 Tense and verbal modality in creating a niche in WPM and percentage of all questions in KIAP-EEFS

Tense	English	%English	French	%French	Spanish	%Spanish
Past	3	11.1	0	0	0	0
Present	13	4.2	23	18.4	7	7.1
Future	3	14.3	0	0	0	0
No tense	13	13.9	8	6.5	12	13.2
Modality	13	16.1	5	7.4	0	0

6.5.5.2 Sentence type

As presented in Table 6-40, the English data uses mainly simple and complex sentence types in questions, which follows overall patterns presented in Section 5.7.3. French data uses mainly complex sentence types in questions and Spanish uses mainly compound-complex sentence types in questions. However, of note is the unusually frequent use of questions that use compound questions in English, where 13% of all such questions in English occur in this function. French similarly follows the overall pattern described in Section 5.7.3 with the notable use of complex questions. Spanish makes most use of compound-complex sentences in this function. This not only differs from overall trends discussed in Section 5.7.3, but this function, which only accounts for 8.2% of all questions found in Spanish, captures 20% of questions composed of compound-complex sentence, as in the following example:

Uno de los aspectos más importantes de este análisis consiste en saber si la intervención política podría eliminar el trabajo infantil mediante leyes que prohíban o impongan la educación obligatoria (specon39)

Table 6-40 Sentence type in creating a niche in WPM and percentage of all questions in KIAP-EEFS

Sentence type	English	%English	French	%French	Spanish	%Spanish
Simple	15	9	8	10.3	5	5
Complex	18	9.3	25	12.9	5	8.3
Compound	8	13	0	0	0	0
Compound-complex	0	0	3	14.3	10	20
No clause	3	12.5	0	0	0	0

6.5.6 Creating a niche: conclusion

For questions that create a niche, the following findings surrounding their most frequent features are of note:

1. English uses short questions, French uses middling length questions, and Spanish uses long questions for this function.
2. English uses mainly direct, Spanish uses mainly indirect, and French uses an equal amount of both direct and indirect questions.
3. English uses mainly direct content questions, French uses indirect content questions, and Spanish uses indirect polar questions.
4. These questions occur mainly in introduction sections.
5. There is no clear use of passivity in questions in this function.
6. Present tense is most common here in each language with English and Spanish using some past tenses.
7. English uses mainly simple and complex sentences, French uses mainly complex, and Spanish uses mainly compound-complex sentences.

When considering questions that create a niche, some distinctive characteristics emerge in each language when compared to other functions:

1. Short questions in English are unusually frequent here. As are mid-length questions in French and long questions in Spanish.
2. Direct questions are unusually frequent in this function in English, while indirect are unusually frequent in French and Spanish.

3. English makes unusually frequent use of declarative questions and indirect content questions. French makes unusually frequent use of direct and indirect content questions, and Spanish makes unusually frequent use of indirect polar and alternative questions.
4. These questions are unusually frequent in introduction sections.
5. English makes unusually frequent use of passives in this function.
6. French follows overall tense and verbal modality patterns presented in Section 5.7.2, while English makes notable use of modality and future tenses.
7. English makes unusually frequent use of compound sentence types and French and Spanish make unusually frequent use of compound-complex sentence types.

6.6 Expressing an attitude and counter-claiming

6.6.1 Expressing an attitude and counter-claiming: frequency

Overall, questions that express attitude and counter-claim occurred rather infrequently in each subcorpus, with French using slightly more, as shown in Table 6-22.

Table 6-41 Questions that express attitude and counter-claim in WPM and percentage of all questions in KIAP-EEFS

	English	French	Spanish
Total in WPM	10	18	5
Percentage %	2	5	2

6.6.2 Expressing an attitude and counter-claiming: sentence length

For English, questions of this function account for 2%, 1.8%, and 2.1% of all 15 or fewer, 16-25, and 26 or more words question lengths, as reported in Table 6-42 and Table 6-43. This renders their distribution within this function relatively normal.

Table 6-42 Expressing attitude and counter-claiming question length statistics in KIAP-EEFS

Sentence length	English	French	Spanish

Average question length in words	33.5	33.3	42.5
Median question length in words	16	27	42.5
Mode question length in words	N/A	22	0

Table 6-43 Expressing attitude and counter-claiming question length in WPM and percentage of all questions in KIAP-EEFS

Sentence length	English	%English	French	%French	Spanish	%Spanish
15 words or fewer	5	2	3	3	0	0
16-25 words	2	1.8	5	3.8	0	0
26 words or more	3	2.1	10	9.3	5	8.7

For French these questions are unusually long. This is determined owing to the fact that for a function that accounts for 5% of all questions in French, it contains 9.3% of all questions that are 26 words or more. For Spanish, like English, there are too few examples to draw any reliable conclusions about this function. However, it is of note that, like French, in a function which accounts for 2% of all questions in Spanish, it contains 8.7% of all questions that are 26 words or more. In the case of English, Table 6-42 presents statistics that appear contradictory where average word length is high, but the median would fall in the middle category. This is due to an outlier in a question that is 94 words long:

Given that we interpret our findings by distinguishing between two ways of redistributing income, a ‘direct’ one, transferring income from the skilled to the unskilled, and an ‘indirect’ one, through manipulations of equilibrium quantities and wages, the reader may wonder why we do not explicitly model these two policy tools separately, for example by allowing the government to supplement a tax scheme with a minimum wage regulation - in the line of several papers modelling the interaction of minimum wage policies and international trade (Brecher, 1974; Krugman, 1995; Davis, 1998; Davis and Reeve, 1998). (engecon05)

Here, the use of “the reader may wonder” opens a question in the readers’ minds which serves to criticise previous approaches.

6.6.3 Expressing an attitude and counter-claiming: question type and form

The types of questions used for this function are illustrated in Table 6-44 and Fig 6-10. English makes similar use of direct and indirect questions where 1.6% of all direct and 2.6% of all indirect question in English occur in this function. French and Spanish make infrequent use of direct questions where 2.7% and 0% of all direct questions in French and Spanish, respectively, occur in this function. Conversely, both languages exhibit greater use of indirect questions, with 9.3% and 6.1% of all indirect questions in French and Spanish, respectively occurring in this function.

Table 6-44 Percentage direct and indirect questions in expressing and argument and counter-claiming in
KIAP-EEFS

Question category	English	French	Spanish
%Direct questions	1.6	2.7	0
%Indirect questions	2.6	9.3	6.1



Figure 6-10 Question types and forms in expressing attitude and counter-claiming in WPM in KIAP-EEFS

English question forms exhibit no real variation from overall trends seen in Section 5.5. For French, most notable in Table 6-45 is the use of indirect polar questions. This function, which accounts for only 5% of all questions in French, contains 31.3% of all the indirect polar questions used in French. Spanish too, albeit to a lesser degree, frequently uses indirect content questions in this function, with 11.1% of all indirect content questions being used to express an attitude and counter-claim.

Table 6-45 Percentage direct and indirect questions in expressing attitude and counter-claiming per type and form in KIAP-EEFS

	Content	Polar	Declarative	Alternative	Elliptical
%English direct questions	2.5	0	0	0	0
%English indirect questions	2.8	3	0	0	0
%French direct questions	4.3	2.8	0	0	0

%French indirect questions	0	31.3	0	0	0
%Spanish indirect questions	11.1	0	0	0	0

6.6.3.1 IFID for questions that express an attitude and counter-claim

With regard to IFIDs it is noteworthy that no nouns or adjectives are used to generate questions that express attitude and counter-claim, as presented in Table 6-46.

Table 6-46 IFIDs that generate questions that express attitude and counter-claim in WPM and percentage of all questions in KIAP-EEFS

IFID	English	%English	French	%French	Spanish	%Spanish
Verb	5	3.7	13	10.6	5	9.1

This function accounts for 3.7%, 10.6%, and 9.1% of all questions raised by verbs in English, French, and Spanish respectively, rendering verbs as an IFID unusually frequent for this function.

In terms of the verbs used to raise indirect questions, the only recurring verb choice is in French, with *demander* as can be seen in Table 6-47.

Table 6-47 Verbs that generate questions that express attitude and counter-claim in WPM and percentage of all questions in KIAP-EEFS

Rank	English		French		Spanish	
1	Explore	3	Demander	5	Precisar	2
2	Wonder	3	Interroger	3	Ahorrar	2
3			Poser	3		
4			Mettre	3		

Questions that express attitude using *demander* act as follows:

On peut se demander si ce modèle est robuste à une situation où les entreprises choisissent simultanément le capital et le prix (frecon44)

Bien que ce point ait été rarement abordé dans la littérature, on peut se demander s'il n'est pas réducteur de ne considérer que la volatilité prévalant entre les pays étudiés pour expliquer l'évolution de leurs flux commerciaux (frecon04)

In both the examples above, *demandeur* is used to criticise models and approaches in the field. Though not frequent enough to be considered a trait of such questions, the English data uses “wonder” in a similar way:

[...] the reader may wonder why we do not explicitly model these two policy tools separately [...] (engecon05)

Spanish presents little in the way of interesting verb use in this function. The few examples that exist, use verbs like *precisar* to critique the work of others, as in the following indirect question:

Las normas actualmente se limitan a señalar la realización de muestreos representativos de las aguas descargadas, pero no precisan con qué grado de confianza deberán establecerse ni reportarse los resultados de tales muestreos (specon21)

6.6.4 Expressing an attitude and counter-claiming: location

In terms of location, there are no clear patterns emerging from the little data available. There is a spread of use across the introduction and middle sections in English and French, and the middle and conclusion sections in French and Spanish, as illustrated in Table 6-48. Of note in English and French is that 4.9% and 7.9% of all questions in introductions, respectively, occur in this function. In Spanish, 50% of all questions in conclusions occur here, although, given the infrequency of questions in conclusion sections in specon, this finding is too small to be of any real circumstance.

Table 6-48 Expressing attitude and counter-claiming question location in WPM and percentage of all questions in KIAP-EEFS

Section	English	%English	French	%French	Spanish	%Spanish
Introduction	5	4.9	8	7.9	0	0
Middle	5	1.5	10	5.1	2.5	1.6
Conclusion	0	0	0	0	2.5	50

6.6.5 Expressing an attitude and counter-claiming: passivity, tense and verbal modality, and sentence type

Given the infrequency of this function there is little data available to identify patterns of use for this function in terms of passivity, tense and verbal modality, or sentence type, as Table 6-49 illustrates. What data there are correspond to overall patterns of use presented in Section 5.7. French and English do not use passives whereas Spanish does. However, despite 7.1% of all passives occurring here in Spanish, this use is still too infrequent to be considered a trait of this function. English and French use mainly modals as 6.5% and 11.1% of questions raised through verbal modality occur in this function. There is also a notable use of past tense for English and French in this function with 11.1% and 33.3% of questions raised through past tenses occurring here. This is complimented by a lack of use of present tense, accounting for 0.8% and 4.1% of present tense questions in English and French. Spanish does not exhibit any use of note in this function, as there are too few questions that perform this function in Spanish to be able to draw any conclusions. Finally, regarding sentence type, no clear patterns emerge in terms of overall frequency. French uses mainly complex questions; however these only account for 5.6% of all complex questions in French. What occurs unusually frequently here are compound-complex questions where 7.4%, 14.3%, and 10% of all such questions in English, French, and Spanish occur in this function.

Table 6-49 Passivity, tense and verbal modality, and sentence type in expressing and argument and counter-claiming in WPM and percentage of all questions in KIAP-EEFS

Equivalence	English	%English	French	%French	Spanish	%Spanish
Not passive	10	2.5	18	9.1	2	1.4
Passive	0	0	0	0	2	7.1
Unclear	0	0	0	0	0	0
Past tense	3	11.1	5	33.3	0	0
Present tense	3	0.8	5	4.1	2	2.4
Future tense	0	0	0	0	0	0
No tense	0	0	0	0	2	2.6
Modality	5	6.5	8	11.1	0	0

Simple sentence	3	1.5	5	6.9	0	0
Complex sentence	3	1.3	10	5.6	0	0
Compound sentence	0	0	0	0	0	0
Compound-complex sentence	5	7.4	3	14.3	5	10
No clause	0	0	0	0	0	0

6.6.6 Expressing an attitude and counter-claiming: conclusion

The main findings for questions that express an attitude and counter-claim are presented, first in terms of most their frequent characteristics.

1. English questions in this function appear short while French and Spanish questions are long.
2. English uses both direct and indirect questions here while French and Spanish use mainly indirect questions.
3. English uses primarily direct content questions, French uses mainly indirect polar questions, and Spanish makes most use of indirect content questions.
4. These questions occur in introduction and middle sections in English, middle sections in French, and middle and conclusion sections in Spanish.
5. There is no strong use of passivity in these questions.
6. French and English use mainly verbal modality, while Spanish uses present tense.
7. French uses complex sentence types, while English and Spanish use compound-complex sentence types in questions.

While the previous summary focussed on what is most common, the following highlight what is most distinctive within this function:

1. French and Spanish questions are unusually long in this function.
2. Each language makes unusually frequent use of indirect questions.
3. French makes unusually frequent use of indirect polar questions and Spanish makes unusually frequent use of indirect content questions.
4. Introduction sections in English and French make unusually frequent use of this type of question while the same is true of conclusion sections in Spanish.
5. Spanish makes unusually frequent use of passivity.

6. English and French make unusually frequent use of past tenses and verbal modality.
7. Each language makes unusually frequent use of compound-complex sentence types in this function.

6.7 Setting up claims and protecting the writer

6.7.1 Setting up claims and protecting the writer: frequency

This function accounts for 26.3%, 22.5%, and 22.4% of all questions in English, French, and Spanish respectively, as indicated in Table 6-50. This makes this function the second most frequent function in each language.

Table 6-50 Questions that set up claims and protect the writer in WPM and percentage of all questions in KIAP-EEFS

	English	French	Spanish
Total in WPM	133	73	53
Percentage %	26.3	22.5	22.4

6.7.2 Setting up claims and protecting the writer: sentence length

English and Spanish use short questions while French questions are longer, as shown in Tables 6-51 and Table 6-52. However, comparing Table 6-52 to overall trends described in Section 5.4, each language shows that questions of 15 words or less are strongly related to this function, as 30.3%, 33.3%, and 30.8% of all questions of 15 words or fewer in English, French, and Spanish occur in this function.

Table 6-51 Setting up claims and protecting the writer question length statistics in KIAP-EEFS

Sentence length	English	French	Spanish

Average question length in words	14.8	19.28	16.27
Median question length in words	11	20	9
Mode question length in words	10	20	9

Table 6-52 Setting up claims and protecting the writer question length in WPM and percentage of all questions in KIAP-EEFS

Sentence length	English	%English	French	%French	Spanish	%Spanish
15 words or fewer	75	30.3	28	33.3	39	30.8
16-25 words	35	25.5	30	22.6	7	13
26 words or more	23	19.1	15	14	7	13

6.7.3 Setting up claims and protecting the writer: question type and form

The distribution of question types for setting up claims and protecting the writer is illustrated in Table 6-53 and Figure 6-11.

Table 6-53 Percentage direct and indirect questions in setting up claims and protecting the writer in KIAP-EEFS

Question category	English	French	Spanish
Direct questions	32	30.7	24.6
Indirect questions	17.1	11.1	18.2

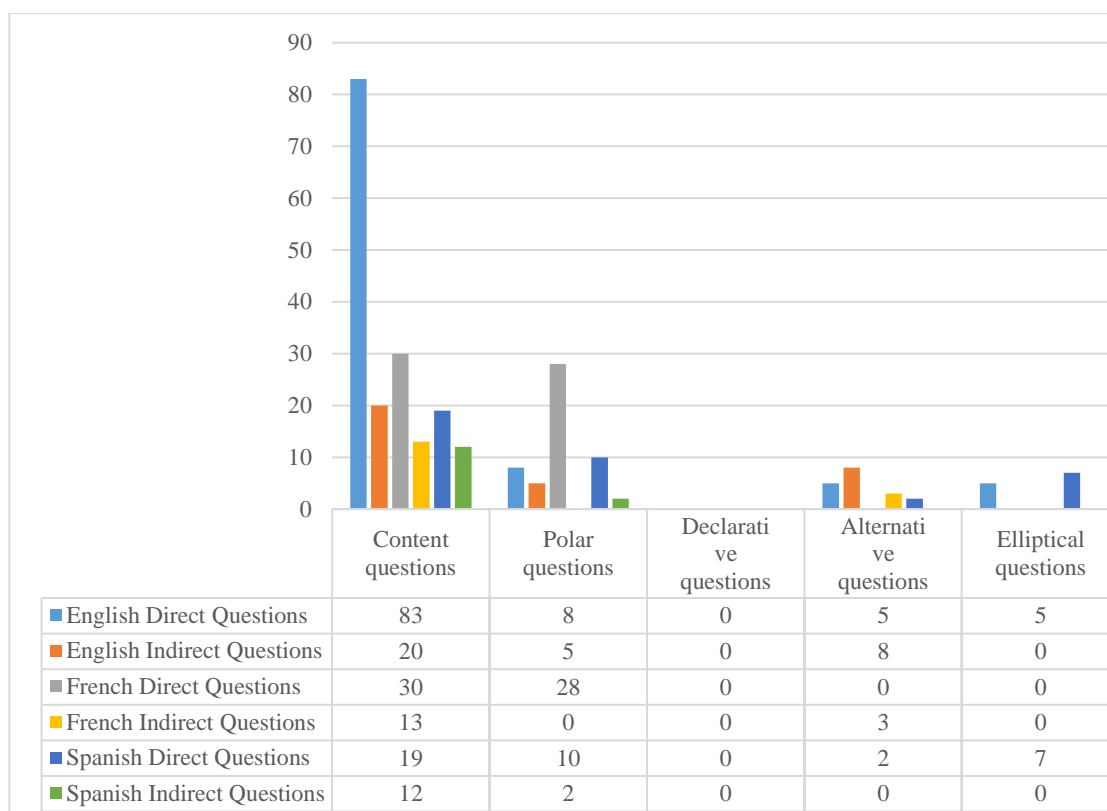


Figure 6-11 Question types and forms in setting up claims and protecting the writer in WPM in KIAP-EEFS

In English, 32%, in French, 30.7%, and in Spanish, 24.6% of all direct questions occur in this function. Similarly, indirect questions are unusually infrequent, as only 17.1%, 11.1% and 18.2% of all indirect questions in English, French, and Spanish, respectively, occur in this function. This is particularly noteworthy given that this function captures 26.3%, 22.5%, and 22.4% of all questions in English, French, and Spanish.

English uses mainly direct content questions and French and Spanish use both direct and indirect content questions and direct polar questions. English, French and Spanish all show evidence of alternative and elliptical questions, albeit to a much lesser degree than content and polar questions. When compared to the overall trends discussed in Section 5.5, it emerges that English, French, and Spanish make a notable use of direct questions, for which relative percentages are reported in Table 6-54.

Table 6-54 Percentage direct and indirect questions setting up claims and protecting the writer per type and form in KIAP-EEFS

	Content	Polar	Alternative	Elliptical
%English direct questions	41.3	10.3	22.2	40
%English indirect questions	22.2	6.1	42.9	0
%French direct questions	52.2	30.6	0	0
%French indirect questions	14.7	0	25	0
%Spanish direct questions	19	33.3	25	42.9
%Spanish indirect questions	27.8	9.1	0	0

Looking deeper into question forms, English makes notable use of direct content questions which account for 41.3% and of all direct content questions in English. This function also uses 40% of all direct elliptical questions in English. However, these are relatively infrequent overall. These questions set up claims and protect writers by posing questions which are immediately answered by the author(s) allowing them to appear knowledgeable and as sources of interest *e.g.*:

What about good [the above finding]? (engecon36)

While setting up claims and protecting the writer only includes 17.1% of all indirect questions in English, the use of indirect alternative questions accounts for 42% of all examples of these questions in engecon. Although these are relatively infrequent overall they are notably used here *e.g.*:

Do trade volumes matter or do they not? (engecon10)

This question is answered directly after it is posed. Its purpose is to raise a doubt in the reader's mind, a doubt which the author addresses.

For French, direct content and polar questions are of particular interest as they account for 52.2% and 30.6% of all direct content and polar questions in French, rendering them of notable use in this function.

Quelle est la fonction objectif que concrétise la politique d'enveloppe globale? (frecon35)

Ne pourrait-on pas considérer la rationalité substantielle comme une forme particulière de rationalité procédurale [Béjean S., Midy F., Peyron C. (1999)] et en conséquence le

réseau d'allocation de ressources comme un cas particulier du réseau de création de ressources? (frecon03)

In Spanish, direct polar and elliptical questions feature particularly highly in comparison to other functions. Direct polar questions and elliptical questions account for 33.3% and 42.9% of all the direct polar questions and elliptical questions respectively, in Spanish. These questions function to set up claims and protect the writer in much the same way as other mentioned throughout this section. That is, by posing and answering questions:

¿Sería eso aconsejable luego de la fuerte crisis macroeconómica? (specon06)

¿Por qué? (specon10)

The latter *¿Por qué?* is preceded by the following sentence and following the *¿Por qué?* is an explanation as to why this is the case.

Así, de forma clara se inscriben en este renglón las competencias derivadas de la política monetaria, y poco más.

Spanish also makes notable use of indirect content questions, which account for 27.8% of all indirect content questions in Spanish.

6.7.3.1 IFID for indirect questions that set up claims and protect the writer

Overall, there is evidence of nouns, adjectives, and verbs being used to raise questions that set up claims and protect the writer as can be seen in Table 6-55.

Table 6-55 IFIDs that generate questions that set up claims and protect the writer in WPM and percentage of all questions in KIAP-EEFS

Parts of speech	English	%English	French	%French	Spanish	%Spanish
Noun	3	6.3	8	50	5	30
Adjective	3	16.7	0	0	2	0
Verb	28	20.4	8	6.4	7	13.6

6.7.3.2 IFID: nouns and adjectives in indirect questions

Of the lesser used IFIDs, there is an unusually infrequent use in English of nouns and adjectives. This is surprising, given that this is the second most frequent question function. That is to say, only 6.3% of all nouns and 16.7% of all adjectives used to raise questions in English occur here. Conversely, although very infrequent overall, French and

Spanish questions that set up claims and protect the writer include 50% and 30% of questions raised by nouns in each language respectively. For French, the only noun used is “question”, which is used in questions such as:

Dans cette optique, la question fondamentale posée aux individus est de savoir combien ils sont prêts à payer pour réduire leur risque. (frecon08)

Spanish also uses *pregunta* as well as other nouns such as *ambigüedad*. in a similar way as can be seen in the following example:

La pregunta que surge forzosamente es cómo se explican estas relaciones entre desigualdad e inversión (specon48)

Hay ambigüedad acerca de cuál categoría del gasto sea favorable para la mayor tasa de crecimiento a largo plazo (specon49)

6.7.3.3 IFID: verbs in indirect questions

English makes use of verbs while French and Spanish appear not to. However, overall, verbs used to raise questions in this function are unusually infrequent, with only 20.4%, 6.4%, and 13.6% of all verbs used to raise indirect questions in English, French, and Spanish occurring in this function. Looking at specific verbs, 50% of the uses of “ask” to raise questions occur in this function. These questions set up claims and protect the writer by proposing solutions to established questions and doubts which allow the writer to make tentative claims more strongly:

As a second preliminary test we ask whether breaking the sterling link slowed the growth in the openness of the Irish economy. (engecon28)

For French and Spanish however, there are no clear patterns of use in verbs for this function.

6.7.4 Setting up claims and protecting the writer: location

Regarding location, Table 6-56 shows that questions used to set up claims and protect the writer occur mainly within the middle section in each language. For English it is of note that this function accounts for 29.3% of questions in middle sections and 27.8% of questions in conclusion sections. For French, this function includes 25.3% of all questions in middle sections and 50% of questions in the conclusion sections. For both English and French, however, it should be remembered that for Spanish, questions do not occur very often in conclusion sections. For Spanish only the middle section is of note, as this

function, which captures 22.4% of all questions in Spanish, contains 31.7% of all questions that occur in middle sections in specon.

Table 6-56 Setting up claims and protecting the writer question location in WPM and percentage of all questions in KIAP-EEFS

Section	English	%English	French	%French	Spanish	%Spanish
Introduction	23	22	13	13.2	5	6.9
Middle	98	29.3	50	25.3	58	31.7
Conclusion	13	27.8	10	50	0	0

6.7.5 Setting up claims and protecting the writer: passivity, tense and verbal modality, and sentence type

This function presents a relatively even distribution of passivity and non-passivity that reflects the overall findings in Section 5.7.1. That being said, this function accounts for 28.6% of all passive questions in French and Spanish, as presented in Table 6-57. This is notable given that this function accounts for 22.5% and 22.4% of all questions in French and Spanish, respectively. Spanish contains many unclear examples, likely owing to the use of elliptical questions. In fact 38.5% of all unclear sentences occur in this function in Spanish.

Table 6-57 Passivity in setting up claims and protecting the writer in WPM and percentage of all questions in KIAP-EEFS

Passivity	English	%English	French	%French	Spanish	%Spanish
Not passive	106	27.1	53	27.3	32	18.3
Passive	8	25	5	28.6	10	28.6
Unclear	20	23.5	15	13.3	12	38.5

In terms of tense and verbal modality, Table 6-58 and Figure 6-12 shows that present tense, no tense, and modality feature strongly for each language.

Table 6-58 Percentage tense and verbal modality in setting up claims and protecting the writer in KIAP-EEFS

Tense	English	French	Spanish
Past	11.1	0	0
Present	28.1	24.5	14.3
Future	28.6	0	50
No tense	22.2	13	21.1
Modality	25.8	40.7	53.8

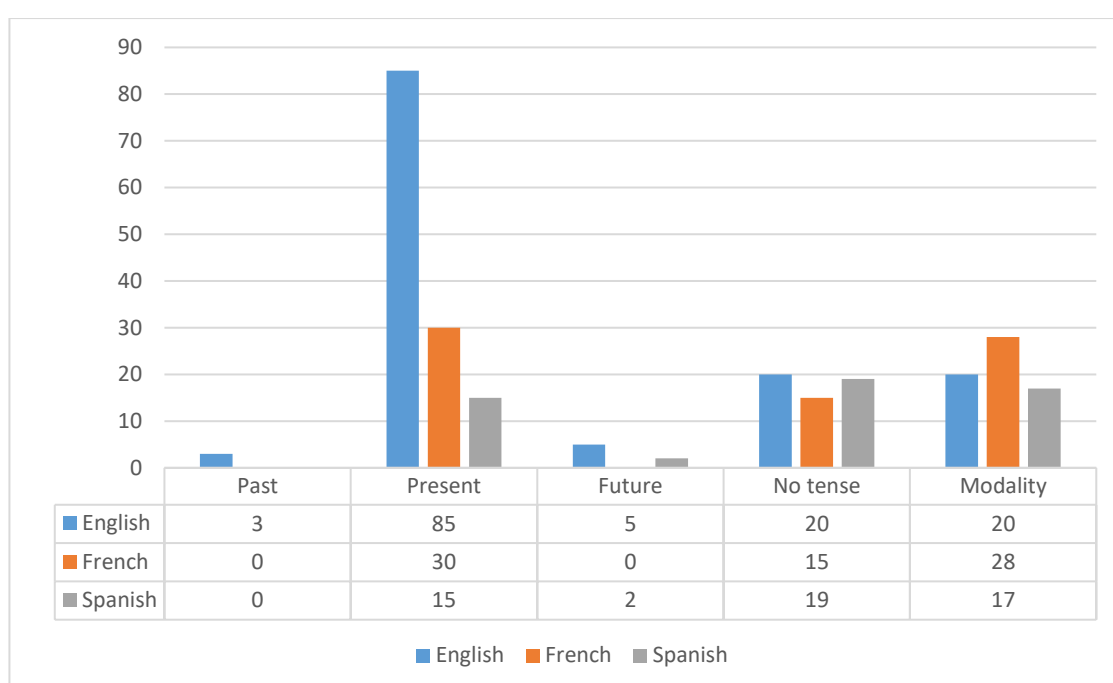


Figure 6-12 Tense and verbal modality in setting up claims and protecting the writer in WPM in KIAP-EEFS

For English, present tense features unusually frequently here with 28.1% of all present tense questions occurring in this function.

Why then are the effects of FPRs on exports weak? (engecon04)

Most interesting for French and Spanish is modality where 40.7% of French questions and 53.8% of Spanish questions that use verbal modality occur in this function.

Le médecin référent peut-il devenir le chef de file d'une organisation en réseau (hiérarchisée) du système de santé? (frecon03)

¿En qué situación se hallaría entonces el reino? (specon03)

The final criteria considered are those presented in Table 6-59 and Figure 6-13 which show simple sentences are common for all three languages, and French sentences use many complex questions.

Table 6-59 Percentage sentence type in setting up claims in KIAP-EEFS

Sentence type	English	French	Spanish
Simple	38.8	48.3	25
Complex	14.7	18.1	12.5
Compound	26.1	22.2	20
Compound-complex	29.6	0	20
No clause	25	0	75

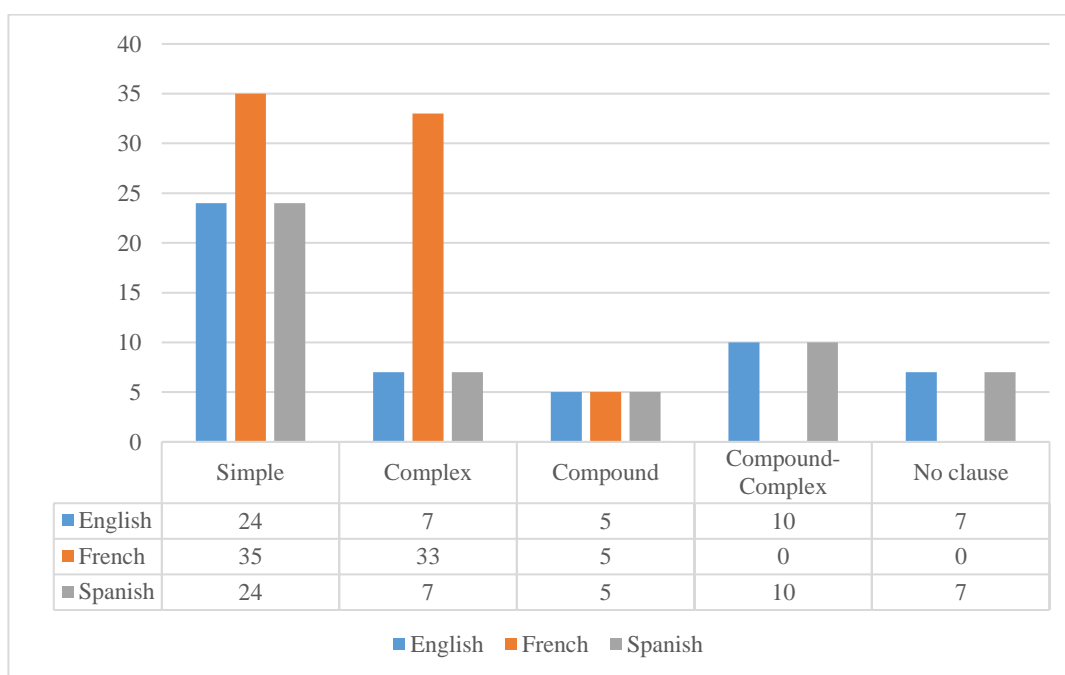


Figure 6-13 Sentence types in setting up claims and protecting the writer in WPM in KIAP-EEFS

When compared to Section 5.7.3, simple sentence structures feature more frequently in this function than others where 38.8%, 48.3% and 25% of simple sentence types in English, French, and Spanish respectively occur within this function. In English,

compound-complex questions also feature highly, with 29.6% of all questions containing compound-complex clause structures occurring within this function.

6.7.6 Setting up claims and protecting the writer: conclusion

For the questions used to set up claims, the following key insights are extracted in terms of most frequent patterns of use:

1. English and Spanish use short questions for this function while French questions are long.
2. All languages use direct questions mainly.
3. All languages use mainly direct content questions.
4. Each language mainly uses these questions in the middle sections.
5. Passivity is not noticeably common in any language for this function.
6. English uses mainly past tenses, French uses mainly past tenses and modal verbs, and Spanish uses mainly modal verbs.
7. All three languages use mainly simple sentence structures for this function.

When considering questions that set up claims, some distinctive characteristics emerge in each language when compared to other functions:

1. Each language uses an unusually frequent number of short questions in this function.
2. Direct questions are unusually frequent here.
3. English uses an unusually frequent amount of direct content, indirect alternative and direct elliptical questions. Direct content questions appear salient for French here, as do direct polar questions and indirect alternative questions. For Spanish direct polar, indirect content, and direct elliptical questions are unusually frequent here.
4. Question generated claims are unusually frequent in middle and conclusion sections in English and French and in conclusion section in Spanish.
5. Passivity is unusually frequent in all languages in this function.
6. For tense, English makes unusually frequent use of present and future tense questions and French and Spanish make notable use of modality. Spanish also makes unusually frequent use of future tenses.
7. For sentence complexity simple sentences feature highly for this function for each language. English also makes unusually frequent use of compound-complex sentences.

6.8 Asking real questions

6.8.1 Asking real questions: frequency

This function is more common than questions that get attention and focus the reader, and express an attitude or counter-claim. However, overall it is rather infrequent, accounting only for 6.9%, 6.2% and 5.1% of questions in English, French and Spanish respectively, presented in Table 6-60.

Table 6-60 Questions that ask real questions and point forward in WPM and percentage of all questions in KIAP-EEFS

	English	French	Spanish
Total in WPM	35	20	12
Percentage %	6.9	6.2	5.1

6.8.2 Asking real questions: sentence length

These questions follow the overall trends of question length presented in Section 5.4. That is to say, the data presented in Table 6-61 and 6-62 show that English real questions are shorter than Spanish and both English and Spanish are slightly shorter than French. The only notable divergence from the overall pattern of use Section 5.4 is that in the case of real questions, shorter questions are unusually frequent in English where 8.1% of all short questions occur here. For French 11.6% of all longer questions occur here and for Spanish, 17.4% of all middle length questions occur here.

Table 6-61 Real questions question length statistics in KIAP-EEFS

Sentence length	English	French	Spanish
Average question length in words	16.9	27.9	25.8
Median question length in words	14	26.5	21
Mode question length in words	14	N/A	21

Table 6-62 Real questions question length in WPM and percentage of all questions in KIAP-EEFS

Sentence length	English	%English	French	%French	Spanish	%Spanish

15 words or fewer	20	8.1	3	3	2	1.9
16-25 words	10	7.3	5	3.8	10	17.4
26 words or more	5	4.3	12	11.6	0	0

6.8.3 Asking real questions: question type and form

Indirect questions are most common in each language with 10.5%, 13%, and 15.2% of all indirect questions in English, French, and Spanish occurring in this function, as presented in Table 6-63 and Figure 6-14.

Table 6-63 Percentage direct and indirect questions in real questions in KIAP-EEFS

	English	French	Spanish
%Direct questions	4.8	1.3	0
%Indirect questions	10.5	13	15.2

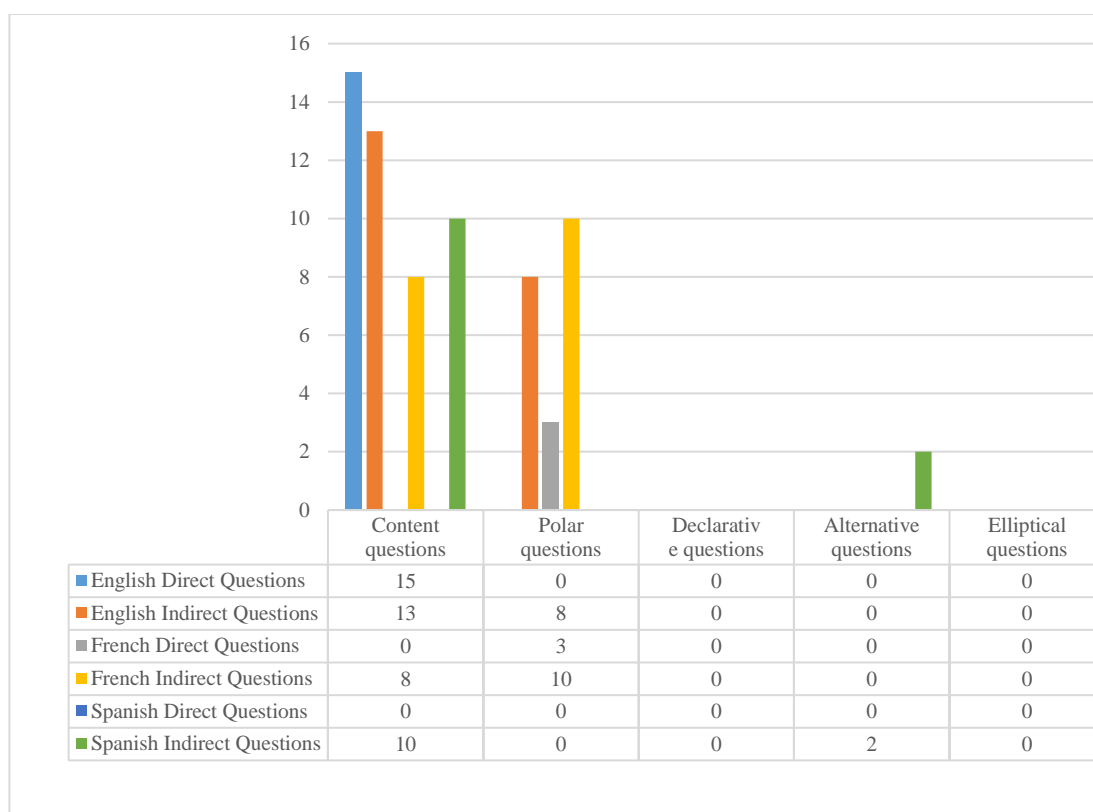


Figure 6-14 Question types and forms in real questions in WPM in KIAP-EEFS

While direct questions are unusually frequent overall, English does make notable use of direct content questions. In fact, 7.5% of all direct content questions occur in this function, which is reported in Table 6-64. More noteworthy however, is the use of indirect content and polar questions in English, which account for 13.9% and 9.1% of all the indirect content and polar questions in the English subcorpus. With these question forms, the authors raise questions that they cannot answer within the text. This positions the author as a problem finder, pointing to questions that are beyond the scope of the paper. The following examples show direct content, indirect content, and indirect polar real questions in English fulfilling this function:

How can the situation be improved? (engecon15)

We have also not explored why different country-decades have different sector biases of SBTC. (engecon23)

A key question for future work is whether incorporating these features into the nominal rigidities model would allow to capture simultaneously the high volatility of exchange rates and the other key macroeconomic facts considered here. (engecon01)

Table 6-64 Percentage direct and indirect questions in real questions per type and form in KIAP-EEFS

	Content	Polar	Alternative
%English direct questions	7.5	0	0
%English indirect questions	13.9	9.1	0
%French direct questions	0	2.8	0
%French Indirect questions	8.8	25	0
%Spanish indirect questions	22.2	0	25

For French, indirect content and polar questions in this function are noteworthy as they account for 8.8% and 25% of all indirect content and polar questions in frecon. Such questions in French are often posed as questions that are beyond the scope of the project, for example:

Au plan théorique, voire épistémologique, on peut s'interroger sur la compatibilité entre ces deux perspectives, entre autres en raison des deux types de rationalité qu'elles présupposent (frecon03)

On n'aurait bien évidemment pas à leur poser la question s'ils supportaient eux-mêmes globalement leur propre risque puisque leurs choix effectifs de prévention donneraient automatiquement la réponse (frecon08)

In the case of Spanish, there are no direct questions. Therefore, the absence here is noteworthy. In lieu of direct questions, indirect content and alternative questions feature for real questions. These account for 22.2% of all indirect content questions, 25% of all indirect alternative questions in Spanish. Given this function only contains 5.1% of all the Spanish questions identified, these are particularly frequent question forms, rendering them clear characteristics of real questions in Spanish. The following examples show how these indirect questions function to point to larger questions about society and obscure questions with no real answers.

Las crisis cambiarias ocurridas durante los últimos años han abierto de nuevo la cuestión de cuál es el sistema cambiario apropiado para las economías en vías de desarrollo y emergentes (specon16)

Es difícil saber si ha sido un error estratégico o una cesión inevitable ante el poderoso empuje de la nueva Alemania (specon15)

6.8.3.1 IFID for indirect questions that ask real questions

For IFIDs for real questions, Table 6-65 shows that verbs and nouns occur to raise questions.

Table 6-65 IFIDs that generate questions that ask real questions in WPM and percentage of all questions in KIAP-EEFS

IFID	English	%English	French	%French	Spanish	%Spanish
Noun	8	18.8	3	16.7	5	20
Adjective	0	0	0	0	2	100
Verb	13	9.3	15	12.8	5	9.1

6.8.3.2 IFID: nouns and adjectives in indirect questions

Overall, nouns used to raise indirect questions occur relatively frequently in this function with 18.8% of nouns English, 16.7% of nouns in French, and 20% of nouns in Spanish occurring within this relatively infrequent function. As such, these IFIDs are unusually frequent in this function. The same is said for Spanish adjectives of which 100% occur in this function; however this is negligible given the infrequent use of adjectives as question IFIDs in Spanish overall, as reported in Section 5.5.3.

Table 6-66 Nouns and adjectives that generate questions that ask real questions in WPM in KIAP-EEFS

Rank	English		French		Spanish	
1	Question	8	Question	3	Cuestión	2
2					Cuestionable	2

Following Table 6-66, each language uses the same noun: “question” and Spanish uses the adjective *cuestionable* which is the only example of an adjective IFID in specon. The following examples show these IFIDs asking real questions that cannot be answered within the texts and point out to future questions that will be in need of answers:

Also left for future research are questions as to the efficiency advantages that may come from networks or long-term buyer–supplier relationships. (engecon39)

On n'aurait bien évidemment pas à leur poser la question s'ils supportaient eux-mêmes globalement leur propre risque puisque leurs choix effectifs de prévention donneraient automatiquement la réponse. (frecon08)

Resulta cuestionable la medida en que esta recomendación mejore efectivamente el aprovechamiento de los estudiantes. (specon44)

Las crisis cambiarias ocurridas durante los últimos años han abierto de nuevo la cuestión de cuál es el sistema cambiario apropiado para las economías en vías de desarrollo y emergentes. (specon16)

6.8.3.3 IFID: verbs in indirect questions

Regarding the verbs used, this function accounts for 9.3%, 12.8%, and 9.1% of all verbs used to raise questions in English, French, and Spanish, respectively. French makes notable use of *interroger* which has a frequency of 10 WPM and accounts for 37% of the use of *interroger* to ask indirect questions. This then appears important for raising questions in this function and is evoked in questions such as:

Au plan théorique, voire épistémologique, on peut s'interroger sur la compatibilité entre ces deux perspectives, entre autres en raison des deux types de rationalité qu'elles présupposent (frecon03)

There are no other patterns in individual verb use across languages.

6.8.4 Asking real questions: location

In terms of location, Table 6-67 shows that real questions occur throughout the research article.

Table 6-67 Real question location in WPM and percentage of all questions in KIAP-EEFS

Section	English	%English	French	%French	Spanish	%Spanish
Introduction	3	2.4	5	5.3	5	6.9
Middle	15	4.5	8	3.8	5	3.2
Conclusion	18	38.9	8	37	2	50

In English, middle and conclusion sections contain examples of real questions. However, of note is the use of questions in conclusions. 38.9% of all questions in conclusions occur in this function, which links the rhetorical use of real questions strongly to conclusion sections in English. French behaves similarly, with 37.5% of all questions in conclusions occurring in this function. For Spanish however, there are too few examples to identify any real trend for locations.

6.8.5 Asking real questions: passivity, tense and verbal modality, and sentence type

The final set of criteria to be considered here are passivity, tense and verbal modality, and sentence type. In terms of passivity, Spanish uses no passive structures. However, as Table 6-68 presents, for English and French, this function accounts for 16.7% and 42.9% of all passivity in questions such as:

How can the situation be improved? (engecon15)

Ces séries, une fois estimées, pouvaient-elles être considérées comme des bruits blancs faibles ? (frecon40)

With regard to tense, English and French make notable use of modality, where 9.7% and 14.8% of all questions raised through modal verbs occur in this function, as can be seen in Table 6-68. For Spanish however, there is no pattern that is distinct for overall patterns presented in Section 5.7.

Table 6-68 Passivity and tense and verbal modality in real questions in WPM and percentage of all questions in KIAP-EEFS

Equivalence	English	%English	French	%French	Spanish	%Spanish
Not passive	20	5.2	10	5.2	7	4.2

Passive	5	16.7	8	42.9	0	0
Unclear	10	11.8	3	2.2	5	15.4
Past tense	3	11.1	5	33.3	0	0
Present tense	13	4.2	0	0	0	0
Future tense	3	14.3	0	0	0	0
No tense	10	11.1	5	4.3	12	13.2
Modality	8	9.7	10	14.8	0	0

Finally, in terms of sentence types, Table 6-69 and Figure 6-15 largely corresponds to the overall findings in Section 5.7.3.

Table 6-69 Percentage sentence type in real questions in KIAP-EEFS

KWIC type	English	French	Spanish
Simple	1.5	3.4	0
Complex	14.7	8.3	16.7
Compound	0	0	10
Compound-complex	7.4	14.3	0
No clause	0	0	0

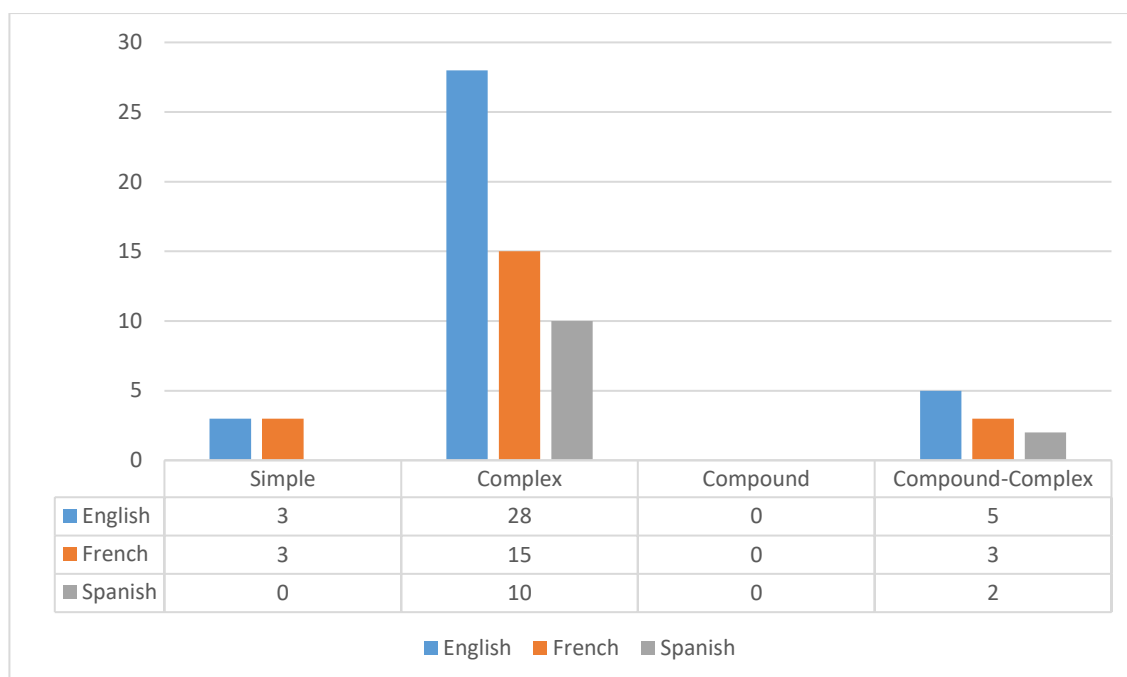


Figure 6-15 Sentence types in real questions in WPM in KIAP-EEFS

One finding of note is the reduction in all sentence types excluding complex sentences. This is particularly noteworthy for English and Spanish for which 14.7% and 16.7% of complex sentences are used in this relatively infrequent function to ask real questions. Therefore, complex sentence types appear to be a characteristic of such questions which is not shared by French where only 8.3% of all complex sentences in questions occur in this function.

6.8.6 Asking real questions: conclusion

Finally, for questions used to ask real questions and point forward, the following insights were extracted, first in terms of most frequent characteristics:

1. English uses short questions, French uses long questions, and Spanish questions are middling in length.
2. All three languages use mainly indirect questions for this function.
3. English uses mainly direct content questions, French uses mainly indirect polar questions, and Spanish uses mainly indirect content questions.
4. English uses these questions in conclusion sections, French in middle sections and conclusion sections, and Spanish in introduction sections and middle sections.
5. Passivity occurs, though not so frequently in English and French. There is no passivity for Spanish.

6. French and English questions are mainly in the present tense or occur with a modal verb.
7. All three languages use mainly complex questions.

When considering real questions, some distinctive characteristics emerge in each language when compared to other functions:

1. English makes unusually frequent use of short questions, French makes unusually frequent use of long questions, and Spanish makes unusually frequent use of questions of middling in length.
2. Each language makes unusually frequent use of indirect questions.
3. English and French make unusually frequent use of indirect content and polar questions, while Spanish makes unusually frequent use of indirect content and alternative questions.
4. Real questions are unusually frequent in conclusion sections.
5. Passivity is unusually frequent in English and French for real questions.
6. English makes unusually frequent use of future tenses, while French makes unusually frequent use of past tenses and modality.
7. English and Spanish make unusually frequent use of complex sentences, while French makes unusually frequent use of compound-complex sentences.

6.9 Conclusion

This chapter has presented the results of the second contrastive analysis of questions in KIAP-EEFS. This section centred on the *tertium comparationis* identified which was the assumption that English, French, and Spanish economics research articles share a rhetorical functional use of question used as reader engagement. By analysing each function, the analysis revealed many similarities and differences in question function both across languages and within them. In Sections 6.2-6.8, each function description is concluded with a presentation of the manifest similarities of question functions across languages in terms of frequent and salient use of equivalences. These show how question functions correspond in English, French, and Spanish. Overall, the results of this chapter and Chapter 5 have offered linguistic descriptions of academic writing in English, French, and Spanish which contribute to descriptions of academic writing and reader engagement in English, French and Spanish, as well as the relationships and relativity between English, French and Spanish academic writing surrounding the use of questions. The following chapter discusses these findings.

Chapter Seven

Questions as reader-engagement in English, French, and Spanish: Discussion of Findings

7 Questions as reader engagement in English, French and Spanish: discussion of findings

7.1 Introduction

This chapter discusses the results presented in Chapter 5 and Chapter 6. In Section 7.2, the focus is on academic writing and reader engagement in English, French, and Spanish, with a specific focus on contribution to the discipline of economics. To address this wider area, the uses and functions of questions as reader engagement are discussed in Sections 7.2.1 and 7.2.2 in order to better understand the differences between English, French, and Spanish. Following that, question length, type and form, location, use of passivity, use of tense and verbal modality, and question sentence type are discussed in terms of the relevant literature presented in Chapter 2. These themes form Sections 7.2.3-7.2.8. Following that, Section 7.3 discusses the contribution this research makes to the multilingual academy and corpus-based contrastive analysis of academic writing. This is followed a brief conclusion in Section 7.4.

7.2 Academic writing and reader engagement in English, French and Spanish

The following discussion focuses on the findings of the analyses presented in Chapters 5 and 6. Reflecting on the identified gaps in the literature, this section pays particular attention to the contribution to research on academic writing in French and Spanish, disciplinary studies of economics research articles, and studies of reader engagement in academic writing.

7.2.1 Reader engagement and questions in English, French, and Spanish

This section aims to contribute to research on reader engagement in English, French, and Spanish academic writing in the economics research article. It focuses on the disparity in the presence of reader engagement markers and moves to explain this disparity in terms of previously published works.

In Section 5.2.1, Table 5-1 presents the frequency of questions identified in KIAP-EEFS used to engage readers in texts. In WPM, these are 505 in English, 323 in French, and 237 in Spanish. Overall English contains the most questions, Spanish the least, and French

sits in between. Moreover, there is a significant difference in the use of questions in English and Spanish, as Figure 5-2 makes clear.

The finding that English uses many questions as engagement markers is not a surprising finding given researchers such as Khedri *et al.* (2015) identify that metatext like self-mentions are more common in disciplines like economics than linguistics. This is further supported by Curry and Chambers (2017) who study questions in linguistics research articles in KIAP. Their study is directly comparable to this study of economics and reveals that English uses many more questions in economics than linguistics. This shows that disciplinary differences occur in the study of metadiscourse, as across the disciplines of economics and linguistics, the frequency of questions is inverted in English and French. Lafuente-Millán's work (2014) too recognises that English makes use of many metadiscourse markers in academic writing. In fact, Hyland (2005b) has described English academic writing as explicit in nature which would explain the quantity of questions present. However, another possible explanation derives from English traditionally being less certain (Hyland 2005a). In this view, questions may reflect Pic and Furmaniak's (2014) identification of questions as being used to signal a lack of certainty. Furthermore, owing to the linear nature of academic writing in English, questions can help organise and emphasise discussions and link ideas, which, again, may explain their frequency in English economics writing (Webber 1994; Dietz 2001; Soler 2007; He and Abdul Rahim 2019).

The disparity in use of questions between English and French is explicable. For example, the findings of this study correspond to Dahl (2004) who finds that metatext occurs less frequently in French economics writing, than in English. That being said, Fløttum *et al.* (2006) find French economics writing to contain rich metatext. This is why, although less apparent than in English, French also contains many examples of questions as reader engagement. With regard to the assertion by Loffler-Laurian (1980) who conceptualises academic French as entirely impersonal and unwilling to risk engagement, this research shows that this is not the case, at least in the case of contemporary economics academic research articles. French uses many questions to interact with readers and most of these are direct in nature. Nonetheless, this study reflects the established view that English

appears to be more explicit in nature than French (Fløttum *et al.* 2006; Carter-Thomas and Chambers 2012).

The lower frequency of question use in Spanish is also somewhat unsurprising. For example, in her study of self-mentions, Mur-Dueñas (2007) finds much less evidence of metadiscourse markers in Spanish, than in English. This is similar to her study of evaluative connectors (Mur-Dueñas 2011), which identifies that Spanish exhibits few examples of evaluative markers. In the same vein, Lorés-Sanz (2011a; 2011b) finds much less use of exclusive pronouns in Spanish, Pérez-Llantada (2010) finds fewer participant-oriented metadiscourse markers in Spanish, and Breeze (2012) considers Spanish as difficult to read. In fact, it has long been argued that metatext is infrequent in Spanish academic writing in general (Valero-Garcés 1996). However, it is not absent, and unlike Lafuente-Millán (2013) who looked at directives and questions in business management research articles, this research finds questions in Spanish to be a valuable rhetorical feature of economics research articles. This view is not shared in Lafuente-Millán (2014) in which questions are found to be infrequent overall, and especially in Spanish. Overall, there is evidence of lower-evaluative marker use in economics (Salas 2015) which goes some way to explain the relatively infrequent use of questions in Spanish, when compared to English and French. However, the overwhelming conclusion to draw from this research on Spanish, is that Spanish uses interactive reader-oriented metadiscourse in the form of questions.

Overall, the use of questions in KIAP-EEFS shows the writers' attempt to hook readers (Ruegg and Sugiyama 2013). Evidently, the importance of this appears reduced in French, and Spanish, at least in terms of questions. There are evident cultural considerations that may explain these differences. It is likely that questions are not considered to be part of the register of academic writing in French, and even more so, Spanish. This reflects Siepmann's (2006) *linguacultures*, where academic registers differ at the cultural level. Given that until very recently, questions were thought to be inappropriate in English academic writing (Swales and Feak 1994), it may be that questions are an emerging feature of academic discourse in French and Spanish. This is a reasonable argument given that changes in use of reader engaging metadiscourse have been documented diachronically within English, with the growth in questions being used over time in

disciplines like biology (Hyland and Jiang 2016). As opposed to language change, language contact may explain the lower or emerging use of questions in French and Spanish. Their emergence in French and Spanish could be owing to the impact of English on other academic languages (Ciapuscio and Otañi 2002). English is often seen as a rule setter for other academic languages (Connor 1996), and traditionally, both French and Spanish, but especially Spanish, have been seen to adopt academic English language norms. This is reflected in research by Salager-Meyer *et al.* (2003). However their study did not explicitly focus on questions. Moreover, given the proliferation of success metrics in publishing being linked to progression, employment, and funding in the academy (Phillipson 2006), lack of questions being used as a rhetorical feature of academic writing in French and Spanish is likely owing to the lack of need to effectively disseminate research in the same ways as English (Ball 2009).

While frequencies do differ, questions occur across all languages in KIAP-EEFS and are normally distributed therein. This widespread use of questions reflects the recognition across all three languages that questions are one of the primary means of overtly interacting with the reader in order to guide their thinking (Thompson 2001a, p.59). In each language, through the use of questions that get attention and ask real questions for example, which will be discussed in the next section, question use reflects Webber's (1994) view of question function *i.e.* questions serve an affective end, creating anticipation, interest, and thought.

Overall, what is clear is that questions as reader engagement markers occur to varying degrees across English, French, and Spanish. These questions differ in frequency, with more in English than French and Spanish, and more in French than Spanish. That being said, they occur across a range of texts and their differences in frequency are understandable as they reflect economics disciplinary factors and English, French, and Spanish cultural factors in academic writing. An important consideration of note is the widely accepted designation of French and Spanish as reader-responsible languages, as discussed in Section 2.2.2, is questionable, based on this research. Relatively, it is reasonable to say that French and Spanish are more reader-responsible than English. However, it would be inaccurate to say that French and Spanish are entirely reader-responsible. Questions as reader engagement are a definitive rhetorical device in each

subcorpus studied in this thesis. Therefore, French and Spanish academic writing is at least somewhat reader-oriented; this topic is considered again shortly in Section 7.3.1. However, next, Section 7.2.2 discusses the findings surrounding the functions of questions.

7.2.2 Function of questions in academic writing in English, French, and Spanish

This section aims to contribute to research on the function of questions as reader engagement in English, French, and Spanish academic writing in the economics research article. The findings surrounding the distribution of functions across and within languages are presented in Section 5.3. Summarising that information, Table 7-1 shows the most and least frequent functions within each language, in descending order. As can be seen, all three languages make almost identical use of functions in order of frequency except for French which inverts the two least frequent functions: expressing and attitude and counter-claiming and getting attention and focussing the reader.

Table 7-1 Ranking of question functions in KIAP-EEFS

English	French	Spanish
Organising the text	Organising the text	Organising the text
Setting up claims and protecting the writer	Setting up claims and protecting the writer	Setting up claims and protecting the writer
Framing the discourse	Framing the discourse	Framing the discourse
Creating a niche	Creating a niche	Creating a niche
Asking real questions	Asking real questions	Asking real questions
Getting attention and focussing the reader	Expressing and attitude and counter-claiming	Getting attention and focussing the reader
Expressing and attitude and counter-claiming	Getting attention and focussing the reader	Expressing and attitude and counter-claiming

With this almost perfect correspondence occurring, the clear message is the similarity in use of question functions to engage readers in each language. It is worth now considering the role of these functions in each language to gain a better understanding of how they were used in academic writing in English, French, and Spanish.

Questions that get attention and focus the reader are most common in English, followed by Spanish, and French. However, overall they are quite infrequent. These questions are constrained to titles and as such are relatively infrequent. However, given the finite space in which they can occur, their relative infrequency is not problematic and they account for 5%, 3%, and 4% of all questions in English, French, and Spanish respectively. Although they are infrequent, the percentage distribution of such questions in English and French is greater in economics than linguistics (Curry and Chambers 2017). This would indicate that this function is more of a feature of economics academic writing than linguistics. In each language, questions are found in titles, and serve to capture the readers' attention (Hyland 2002a) and focus them on the topics of the paper.

Titles are not evidence of ignorance on the part of the writer but are rhetorically used to hook readers (Ruegg and Sugiyama 2013, p.13). Their use corresponds in terms of frequency in each language with slightly fewer examples in French. Interestingly, Alès *et al.* (2016) identify confusion and a lack of clarity in French academic research article titles which has led to their lack of visibility and readership. The lack of visibility and questions in titles is consistent with the Buter and van Raan (2011, p.616) who recognise that the value of questions in titles is to increase readership. The higher use of this question function in English also reflects Cook and Plourde's (2016) argument that questions are used in titles to increase visibility and enticement for readers. Owing to the publish or perish nature of academia, and the pressure to have work cited within the academy (Ball 2009), it is unsurprising that English would see a higher usage of rhetorical devices that increase the chances of dissemination. Given that the current state of publishing culture has derived from English language contexts (Bennett 2015), it makes sense that English would employ more means to access a readership. As the publish or perish culture is not as present in French and Spanish academic writing contexts (Breeze 2015), the presence of questions in titles is not necessarily explained by the phenomenon of accessing increased readership.

Interestingly, Hyland argues that questions can create more dramatic hooks. This may better explain role of questions in titles in French and Spanish. Salager-Meyer *et al.* (2003) identify dramatic writing as more characteristic of French and Spanish academic writing than English. Of note is the fact that this function is slightly more common in

Spanish than French. While the view of Spanish as containing very little metatext is well established (Mur-Dueñas 2007b), it appears that French contains fewer examples for this function. This may be owing to the increased Anglicisation of Spanish academic writing (Ciapuscio and Otañi 2002).

For this corpus study, the function of getting attention is sufficiently broad. However, it is not clear what the motivation is to seek this attention. Is it to be noticed and referenced or is it to create more interesting and exciting texts? Unfortunately, a corpus study cannot answer these questions. It is not possible to access the motivations writers have for choosing to put questions in their titles through the sole analysis of their language. As such, it is not possible to say, in any definitive terms, that questions in titles are employed to increase readership. Nor is it possible to say definitively that in French and Spanish, questions are used in titles to create a more interesting sounding text. However, given the current understanding of the academic cultural traditions of these academic languages, they are reasonable assumptions. As this thesis centres on a corpus study, this chapter regularly draws on wider research on the cultural practices within English, French, and Spanish to propose possible explanations for the findings. Going forward, it would be interesting to interview academics from these different disciplinary and language backgrounds, to identify their motivations and link these to wider functional analyses.

Questions that frame the discourse are most common in English, followed by Spanish and then French. Looking at their proportional distribution within languages, this function accounts for 12%, 12%, and 17% of all questions in English, French, and Spanish respectively. When compared to Curry and Chambers (2017), this function behaves similarly in English linguistics papers but is less frequent in French economics research articles than linguistics, as 19% of questions in French linguistics research papers serve to frame the discourse. Nonetheless, this is the third most common function for each language, and broadly serves to provide a framework for the discourse which is signalled by the posing of research questions (Hyland 2002a). This means that research questions are a common feature of the economics academic research article in English, French, and Spanish.

Spanish makes particular use of this function, which, initially, may appear surprising. Spanish, as discussed in Section 2.7, is recognised as complex, digressive, unclear, and

not reader-friendly (Bennett and Muresan 2016). Yet, Spanish economics articles appear to make proportionally higher usage of questions that frame the discourse. This function serves valuable reader signalling roles which, initially, may call into question held views of Spanish academic writing as not reader-oriented. However, it is possible that the use of such questions is a reflection of what Mur-Dueñas (2011) finds in her study of topicalisers. She finds that the lack of linearity in Spanish necessitated the use of topicalisers to overcome complex argumentation norms. This means that questions that frame the discourse may be used more often in Spanish to account for reader-unfriendliness elsewhere. Nonetheless, the attention paid to the reader to address problems of complexity is very much characteristic of writer-responsible languages which contradicts research that typically sees Spanish as reader-responsible. Overall, these questions, common in Spanish, reflect Soler (2007) as they are used to open discussions and focus on specific areas of interest and the presence of such questions occurring frequently in Spanish is reflected in Lafuente-Millán (2014, p.218) who finds that questions:

tended to cluster at the beginning of the research article and fulfilled at least two purposes: a textual function, by providing a guiding path along which the article is structured, and an interpersonal function, by raising the readers' interests and encouraging them to continue reading.

For organising the text, questions are most frequent in English; however unlike the previous functions, French contains the second highest number of questions, which is followed by Spanish. Proportionally, these questions are the most frequent question function used in each language and they account for 39%, 40%, and 41% of all questions in English, French, and Spanish, respectively. These questions occur in subheadings in each language as well as within the research article text proper and their purpose is to organise the content of the article for the reader.

Effectively, questions that organise the text in each language serve to help readers navigate the text and, following Soler's work (2007), these questions are useful for re-opening discussions. This is a popular use of this function in each language. Compared to linguistics in Curry and Chambers (2017), this function is more frequent in economics in English and French and given the recognition that English takes a linear and clearly signalled approach to academic writing (Hyland and Salager-Meyer 2008), this overall high frequency is unsurprising. It is unsurprising as this function is rhetorically linked

clarity and organisation and can work in a similar way to transition markers which Khedri *et al.* (2013) identify as being particularly frequent in English economics research articles. For French and Spanish, their use may be attributed to the goal of counteracting more complex discourse, typical of French and Spanish academic writing (Mur-Dueñas 2011). Nonetheless, the evidence of text organising questions is somewhat surprising in French and Spanish. This is because researchers like Salager-Meyer (2011b) argue that research articles in French and Spanish contain few cohesion devices. Moreover, research elsewhere has positioned French and Spanish as content-oriented (Oliver del Olmo 2004), which means that they are more concerned with quantity of knowledge presented rather than the way knowledge is presented. To the same effect, Spanish research articles have been positioned as not very linear in the organisation of information (Mur-Dueñas 2011). Of course, relatively, French and Spanish make fewer attempts to organise texts using questions. However, the overall use of questions as an organisational device in French and Spanish contradicts the view of French and Spanish academic writing style as indifferent to the manner in which knowledge and information are presented.

The following two question functions are creating a niche and expressing an attitude and counter-claiming. Questions that create a niche are most common in English, followed by French, and Spanish. These questions account for 9%, 11%, and 8% of questions in English, French, and Spanish, respectively. In each language, these questions served to raise questions about the current state of the art and identify areas to which the research presented in the research articles could contribute. Compared to linguistics in Curry and Chambers (2017), economics uses more of these question types proportionally in English and French and overall, this question function reflects an important facet of Swales' (1990) CARS model regarding the identification of a niche, which is evident in each language.

Overall, the importance of creating a niche has been well established as a feature of academic writing in English, French and Spanish (Swales 1990; 2004; Ryvityté 2003; Pérez-Llantada 2010; Lakić 2010; 2015; Carter-Thomas and Chambers 2012). The higher proportional distribution in French can be understood as reflective of a content-oriented culture as the focus on new knowledge is central (Clyne 1994); however this does not

account for the Spanish data which reflect a lower use of metatext (Mur-Dueñas 2007a; 2007b).

The function of expressing an attitude and counter-claiming is the least frequent in the English and Spanish subcorpora and the second least frequent in French. Interestingly, this function is the only one which is not most common in English. It is most frequent in French, followed by English and Spanish; it accounts only for 2%, 5%, and 2% of all questions in English, French, and Spanish and is less frequent in economics than linguistics (Curry and Chambers 2017). This function is used to evaluate others' ideas and, if an unfavourable evaluation occurs, present an alternative proposition. The following questions show evaluative language like "bother", a criticism of approaches in French with *réducteur de ne considérer*, and a question of adequacy in Spanish with *adecuados*:

Why bother with surrogates when we can look at prices directly? (engecon10)

Bien que ce point ait été rarement abordé dans la littérature, on peut se demander s'il n'est pas réducteur de ne considérer que la volatilité prévalant entre les pays étudiés pour expliquer l'évolution de leurs flux commerciaux (frecon4)

Cuál es la tasa óptima de ahorro de los países y, por otro, si los métodos desarrollados son adecuados para valorar la destrucción y la degradación del medio. (specon34)

English and Spanish make little use of this function. For English, this is likely owing to Salager-Meyer *et al.*'s (2003) assignation of English as not very critical as well as the many other rhetorical devices available to writers to express evaluation such as attitude markers (Hyland 2004a). For Spanish, the lower frequency is likely less to do with the use of other rhetorical devices but with the general lack of metatext and evaluative markers in Spanish (Valero-Garcés 1996; Mur-Dueñas 2011). For French, Van Bon and Swales (2007) and Salager-Meyer *et al.* (2003) consider French as more critical, which could explain the use of metadiscourse in French to take a critical perspective. However, Salager-Meyer *et al.* (2003) also makes the same claim for Spanish which is not corroborated in this function. Therefore, the avoidance of metatext in Spanish or its convergence towards English language norms (Ciapuscio and Otañi 2002) seem to outrank the need to express criticism.

The penultimate function discussed here is setting up claims and protecting the writer which is the second most common function used in each language. These are most

common in English, followed by French and Spanish and account for 26%, 23% and 23% of all questions in English, French, and Spanish. These questions create an opportunity for writers to make claims. In comparison to Curry and Chambers (2017), this question function is equal to the distribution in linguistics for English and French, which shows no disciplinary variation in questions that set up claims in English and French economics and linguistics research papers. Across languages there is also very little difference in their distribution, rendering questions that set up claims a feature of each language.

The questions used for this function aimed to help writers create claims jointly with their readers (Hyland 2002a). Occurring in each language, writers set up claims by asking questions to raise doubts. The shared doubt protects them, as the doubt is also belonging to the reader. In English, lack of certainty (Hyland 2005a) can explain why writers may opt for this rhetorical resource. In setting up claims, the question's purpose is to generate doubt so that the writer can answer it and mitigate any possible rejections. Expression of doubt is a common feature of English economics research articles and questions have been seen to be used by authors to minimise face-threatening certainty (Ball 2009). However, for French and Spanish, this is not necessarily the case, as these languages are known for being authoritative (Salager-Meyer 2011b).

In French, it is more likely that these questions allow the author to present their arguments and stance. Following Salager-Meyer (2011b), who argues that French is very authoritative in academic writing and willing to make claims, the frequency of this function is explicable. This function allows the author to make themselves more present and authoritative through demonstrating their knowledge and proposing explanations for problems areas.

For Spanish, Lafuente-Millán (2014) see engagement markers as crucial for the exploitation and anticipation of reader reactions as well as for the guidance of readers' thinking. These functions reflect questions that set up claims which may explain their importance in Spanish. As discussed in Section 2.9, Orta *et al.* (2004) propose that Spanish academic writing tends to show less stance. However, for Hyland, questions that set up claims exemplify writer stance through questioning. As such, this rhetorical feature may be a means to counteract the lack of stance taking in other parts of the research article.

A further consideration for their use in French and Spanish draws from Bennet (2015). French and Spanish are known for more literary writing styles that are dramatic and prose-like. Given the role these questions play in embodying stance and evaluation, it may be that the decision to use questions to create, share, and dispel doubts reflects this more literary academic writing tradition in French and Spanish. Again, this signals the limitations of a corpus study. The categorisation of questions by function are based on the interpretation of the perceived intention and motivation of the writers. Of course, through a corpus study, it cannot be definitively known what exactly those intentions are.

The final question function, real questions, are most common in English, followed by French, and Spanish and proportionally account for 7%, 6% and 5% of all questions in English, French, and Spanish. In each language, these questions reflected Hyland (2002a) who argues that such questions genuinely call for information which does not form part of the authorial text in which the question is raised. Such questions point to further gaps in knowledge. Real questions reflect Webber (1994) who sees a value in questions for creating anticipation, interest and thought. Their frequency in English reflects the tendency of English academic writers to be the identifier of problems (Hyland 2002a). However, in French and Spanish the lower frequency could be explained by the need to appear authoritative (Salager-Meyer 2011b), and the tendency to focus on appearing knowledgeable (Clyne 1994; Cuenca 2003; Oliver del Olmo 2004). This implies that the importance of identifying future problems may not be as integral to research articles in French and Spanish, at least in the case of economics.

As discussed in Section 2.10.1, the seven functions of questions used in this study build on the functions outlined by Hyland (2002a) by enriching the functions with more recent literature. These functions are broad in definition which allowed for effective analysis across languages. However, this discussion has made clear that there is a difference between perceived question function and the motivation for its use. For example, it is possible to identify a question as one that frames the discourse in each language if it is linked the broad research aims that define and frame a paper. However, it is not possible to identify why this question was used by the writer. In many of Hyland's studies of metadiscourse (2002a; 2002b; 2002c; 2005a; 2005b), he also interviews academics to access their motivation for using metadiscourse. This would be an important follow-up

study for this thesis, in order to better understand the question functions in French and Spanish.

What this section has shown is how questions are used to achieve the same functions in academic writing in English, French, and Spanish, albeit to varying extents. Overall, in order to address these subtle differences across languages for each function, various explanations are proposed throughout the section drawing on research on academic writing in English, French, and Spanish. Next, the length of questions and question length per function are discussed in terms of relevant literature on sentence length in English, French, and Spanish academic writing.

7.2.3 Question sentence length in academic writing in English, French, and Spanish

Most questions in English and Spanish contain 15 words or fewer, with the mode question length being 10 words and 7 words respectively, as reported in Section 5.4. These account for 49.3% and 53.1% of all questions in English and Spanish. French questions are longer with most questions falling between 16 and 25 words and with the mode being 16 words. Such questions in French account for 41.1% of all questions.

This overall view of short questions in English is supported by research which finds sentences in English academic writing to be short in general (Van Bonn and Swales 2007; Sawyer *et al.* 2008; Ghivirigă 2012; Gea-Valor *et al.* 2014). These questions reflect what Sawyer *et al.* (2008) identify as good writing practices in English. For Spanish, while research generally shows sentences to be long in nature (LoCastro 2008; Breeze 2012; Pérez and Ramiro 2015), questions in Spanish in this study appear short. This finding corresponds better to Taft (2011) who finds sentence length parity between English and Spanish, perhaps owing to the Spanish convergence towards English language norms (Ciapuscio and Otañi 2002). This calls into question Spanish's classification as content-oriented (Cuenca 2003; Oliver del Olmo 2004) which characteristically uses long sentences. It may be that for certain types of sentences *i.e.* questions, Spanish uses fewer words. However, from this research it is not clear whether the remainder of sentences present in Spanish use long sentences as described by Bennett and Muresan (2016). What is clear is that short questions are characteristic of questions in economics writing. For French, the findings in this study are consistent with the literature on sentence length.

Questions in economics academic writing have been found to be longer (Van Bon and Swales 2007) yet still concise with the mode length of 16 words corresponding to Benichoux *et al.*'s (1985) guidance on French academic writing. This too corresponds to Liddicoat (1997) who finds the average sentence length was 15.97 words in academic French. That being said, the average and median length of questions in French are somewhat higher in this study, with 23.7 words on average and a median of 22 words. This finding reflects Bennett and Muresan's (2016) and Birch-Bécaas' (1996) work which identify long questions as characteristic of French.

Drawing on the results of Chapter 6, question function sees interesting differences occur in terms of question length. For example, with getting attention and focussing the reader, questions are shorter in English and Spanish when compared to French. Overall, questions in this function are unusually short in all three languages. This particular use of short questions in all three languages likely reflects the need for conciseness (Benichoux *et al.* 1985; Sawyer *et al.* 2008; Taft 2011). That being said, this may also be a constraint from academic journals which often limit accepted word counts of titles for editorial purposes.

For framing the discourse, English uses long questions. In fact English question length is similar to French in this function and both are unusually long, while Spanish questions are short. It is curious that English questions are so much longer than usual; however, it is worth noting that many indirect questions are used in English in this function which may link to the notable length of questions here. This is an anecdotal assumption however, as this study did not test question length as a *tertium comparationis* in terms of question type; this could make for an interesting follow-up study.

For questions that organise the text, English and Spanish use many short questions and French uses questions that are middling in length. However, in this function middling length are unusually frequent in French and Spanish, which proportionally means that questions are longer here than in some other functions. Spanish questions being unusually long diverges from overall patterns of use and reflects general views in the literature from Bennett and Muresan (2016), for example. Rhetorically, this function differs from others as it is explicitly concerned with text organisation. Such cohesion devices are typically infrequent in Spanish (Salager-Meyer 2011b), which renders questions that organise the text an interesting reader-responsible feature of Spanish writing.

Next, questions that create a niche are typically and unusually short in English, of middling length in French, and long in Spanish while for expressing an attitude and counter-claiming through questions, English typically uses short questions, while French and Spanish questions are typically and unusually long. While these functions are very infrequent, it is worth noting that an important facet of these functions is their focus on knowledge. Content-oriented cultures typically focus on new knowledge formation. A further feature of such a culture proposed by this thesis is the use of long questions in the formation of knowledge. This proposition is based on the view that it is possible that question functions that reflect cultural actions more familiar to Spanish academic writing are more likely to maintain Spanish language patterns *i.e.* long sentences.

Questions that set up claims and protect the writer are short in English and Spanish, yet long in French. However, each language uses an unusually frequent number of short questions in this function when compared to other functions. This may appear somewhat surprising as the function of this question reflects the authoritative nature of French academic writing but does not reflect the tendency for long and complicated sentences therein. Finally, real questions are typically and unusually short in English, long in French and middling in length in Spanish which signals that brevity is important in English for questions that show gaps in knowledge. For French, such questions are embedded in more complicated sentences, which is supported by Bennet and Muresan (2016).

Overall, this section has shown that the pattern of question length is the use of short questions in English and Spanish and longer questions in French. Moreover, question length is more dynamic than this, as different functions of questions appear to favour different sentence lengths. The next section discusses the types and forms of questions used in this study in terms of relevant literature.

7.2.4 Question types and forms in academic writing in English, French, and Spanish

In Section 5.5, the overall summary of question type and form argued that questions in this study occur most often as direct questions in each language. English uses more of each question type, followed by French and Spanish. However, proportionally, French uses a higher percentage of indirect questions when compared to other languages, followed by English, and Spanish. In terms of frequent forms, English and Spanish use

mainly direct content questions while French uses mainly direct polar questions and for indirect questions, although less frequent, all three languages use mainly content questions. Each language uses a small number of alternative and elliptical questions where alternative questions demonstrate no distinct patterns of use across languages, and elliptical questions are more common in French and Spanish than English. Declarative questions are the most infrequent overall, with no examples in Spanish.

Very few studies of questions in academic writing have made the distinction between direct and indirect questions. Blagojević and Misic-Ilic (2012) for example study academic discourse in English and Serbian and identify embedded/indirect questions as a feature of academic writing. However, they restrict these to questions formed of ‘subordinate nominal clauses of wh- or if/whether type, preceded by clauses containing an illocutionary force marking device, such as a verb (ask, wonder...) or a noun (question, answer...)’ (2012, p.23). They identify indirect questions as a distinct form of questions. Therefore, they do not include questions that do not have a wh- complement like Romero (2005) and do not subcategorise them as content questions, polar questions, *etc.* as Suñer (1993) suggests. In other studies on reader engagement, by Hyland and Jiang (2016) for example, questions are consistently identified as direct questions only. Therefore, this thesis offers a more detailed view of direct as well as indirect question types and their question forms when compared to existing literature. In fact as there are two types of question (direct and indirect), five question forms (content, polar, declarative, alternative, elliptical), and three languages, this accounts for 30 different combinations of question type and form to consider and compare.

Most studies of questions in the literature focus on direct questions like Nagano (2010) who finds that similar question forms to those identified in this study occur in his study of academic writing. His research finds polar questions to be most common in economics academic writing in English. This does not reflect the findings of this thesis, as direct content questions are most common in English; however, Nagano’s (2010) finding surrounding polar questions does mirror French in this thesis, which uses many polar questions. Like this research, his study also finds evidence of content and alternative questions. Biber *et al.* (1999) corresponds to the findings in this thesis as 55% of questions in their study are content questions and 40% are polar questions. They too find evidence

of less frequent questions forms like alternative questions, declarative questions, and tag questions. There is no evidence of tag questions in this study, but the other questions are also infrequent. In Curry and Chambers' (2017) study of linguistics, French contains mainly content, polar, and declarative questions while English contains content and polar questions mainly. In comparison, this study sees mainly polar questions in French, and fewer direct questions than English. Therefore, there are clear differences in formal use of questions across disciplines. The use of direct questions in English is unsurprising owing to its documented explicitness (Fløttum *et al.* 2006). Conversely, the presence of direct questions being greater than indirect questions in French and Spanish is somewhat surprising given these languages' preference for indirectness and negative politeness strategies (Loffler-Laurian 1980; Clyne 1994; Lafuente-Millán 2014). This indicates, once again, a divergence from reader-responsible and content-oriented language practices in French and Spanish. For French, the greater proportional use of indirect questions can be linked to the French language's preference for negative politeness strategies, while for Spanish, the convergence towards English language norms may account for the increased use of direct questions (Ciapuscio and Otañi 2002).

Content questions necessitate some new information, while polar questions elicit a reaction to given information. As a label, it might seem more fitting that content questions would belong to content-oriented culture. However, this may be a false equivalence. Content questions require content as an answer. As such, the question itself is a metadiscoursal device that helps organise that content. By asking a content question, it allows for the content to be discussed, thus it can serve organisation purposes. Conversely, polar questions appear to be more concerned with embedding the information within the question, necessitating a yes or no response. Therefore, there appear to be clear links between the use of content questions in formal-oriented languages and polar questions in content-oriented languages, based on their definitions and values detailed by Clyne (1994). English and Spanish appear more concerned with delivering knowledge than French, and French seems more concerned with reacting to given knowledge. For Spanish, the content-oriented nature of the language might predict the use of polar questions so this is somewhat surprising (Cuenca 2003; Oliver del Olmo 2004). Curry and Chambers' (2017) indicate that there are also disciplinary implications that could

account for this. That is because content questions are more common than polar in academic writing in French linguistics research articles.

For the smaller questions forms (declarative, alternative, elliptical), there is little of note at this general level other than they reflect other studies which also find these phenomena to be infrequent (Biber *et al.* 1999). Overall, there are evident differences across languages in the use of question types and forms. This difference is increasingly evident at the level of function, as certain question types and forms are used more or less often in different functions and languages.

For getting attention and focussing readers, only direct questions are used. The use of direct questions is unsurprising as questions in titles are seen to ‘hook’ readers (Ruegg and Sugiyama 2013); as such, the question mark is integral to constructing this eye-catching question. Nagano (2010) finds that questions in titles are typically full sentences and appear in the first unit of a two unit title most often. However, this is not the case in this study as elliptical questions are quite common and generate grammatically incomplete clauses.

Across languages, questions that frame the discourse allow the writer to frame the answer to research questions as the presentation of new knowledge when using direct content questions in French and Spanish. In the case of polar questions, writers move towards a specific answer to a question, indirectly in English and Spanish, and directly in French. Overall, these questions are typically used to raise research questions in the research article *i.e.* questions that have guided the study presented therein. As such the polarity reflects the case that in each language there are expectations to answer a research question with either yes or no. Curry and Chambers (2017) find polar questions to be most common for this function in linguistics in English and French which means that economics and linguistics research articles in English and French use questions similarly. For French and Spanish, the use of polarity reflects the focus that research questions tend to necessitate a yes or no answer. The indirect nature again is not surprising given their description as content-oriented which reflects an indirect approach to communication (Cuenca 2003; Oliver del Olmo 2004). The use of indirect questions in English creates an implicitness that is uncharacteristic of English academic language (Fløttum *et al.* 2006). Given that these questions are linked to the raising of research questions that a research paper

addresses, it is reasonable to assume that the choice to use such a question type and form is not coincidental but a rhetorical decision by the writer. One possible explanation for indirect question use may be linked to attempts to mask or hedge questions, which is a common behaviour in English academic writing, specifically in economics (Ball 2009). Indirect questions may be used to manage expectations of the reader as direct questions would be more likely to predict a definitive answer, as discussed in Section 2.10.1. This is an important finding as indirect questions in English are most used in this function.

For organising the text, the use of content questions in each language means that organisation questions are used to structure content or knowledge, which may reflect content-oriented cultures (Clyne 1994). In contrast, Curry and Chambers (2017) find polar questions to be most common for this function in linguistics in English and French. However, in Curry and Chambers (2017), consideration was not given to indirect questions which may account for this difference, as French uses mainly indirect content questions. French's use of indirect content questions reflects its less explicit nature (Fløttum *et al.* 2006). The use of content questions in both Spanish and French reflects the focus on presenting knowledge (Bennett and Muresan 2016). This is because questions used to organise texts in French and Spanish are organising knowledge/content. In English, linearity is more important and as such, the salient use of direct and indirect polar questions is noteworthy as they allow for the information to be presented and discussed, reflecting formal-oriented cultures (Clyne 1994). This view contradicts an earlier proposition on content and polar questions in framing the discourse in this section. The difference in this argument relies on the focus on organising information and knowledge, as opposed to framing the information and knowledge delivered by the research paper. The use of elliptical in French is likely owing to their use in internal subtitles.

Creating a niche in English focuses on the contribution of knowledge, hence the use of direct content questions, rendering it more akin to content-oriented cultures (Clyne 1994). For English, the unusually frequent use of indirect content questions act similarly to those that frame the discourse as their use may be an effort to manage expectations regarding the contribution of the research. For French, the use of content questions, both direct and indirect is interesting. The use of content questions in French signals the importance of

the information or knowledge to be presented, which is characteristic of French (Bennett and Muresan 2016). Interestingly, for creating a niche in Spanish, the use of direct and indirect polar questions places the content in the question and the focus on the validity of the gap identified through necessitating a yes or no answer, or a choice in an alternative question.

For questions forms, direct content questions are most common in English allowing the writer to use the question to elaborate their view of the criticism presented. This privileging of information and knowledge is typical in these data but inconsistent with the view of English as a formal-oriented language. Similarly, the use of indirect questions in French and Spanish reflects their position in the literature as indirect, less explicit, and content-oriented (Clyne 1994; Ciapuscio and Otañi 2002; Cuenca 2003; Oliver del Olmo 2004; Fløttum *et al.* 2006; Bennett and Muresan 2016). Unlike creating a niche, questions that express attitudes in French seem to place the knowledge in the question through polar questions. The questioning of others' work anticipates an agreement based on the constrained choice of yes or no from the reader. For Spanish, content questions seems to allow the writer to use the criticism to discuss their view of the topics.

For setting up claims and protecting the writer, the use of direct content questions in each language allows the claim to focus on knowledge which is consistent with its content-oriented nature of French and Spanish (Bennett and Muresan 2016). However, as in many of the cases presented so far, the importance of new knowledge in English writing does not correspond to its typical formal-oriented nature (Clyne 1994). That being said, English also exhibits alternative and elliptical questions which allow for the evaluation of specified claims, as opposed to their discovering. This is more in keeping with the formal-oriented nature of English. Conversely, French and Spanish's varied use of types and forms shows relative inconsistencies in this function, typical in academic writing in these languages (Fløttum *et al.* 2006; Bennett and Muresan 2016).

For real questions, indirect questions are most common and unusually frequent in each language. This is unsurprising as this is the one question function that does originate in a degree of ignorance, as the writers do not have an answer for these questions. Therefore, the use of indirect questions here can minimise the signalling of a gap in knowledge which can be face-threatening for the writers (Lafuente-Millán 2014). English uses mainly direct

content questions, indirect content, and indirect polar questions all of which are unusually frequent. This shows the range in answers anticipated from these questions in English. French uses mainly indirect polar questions which are also unusually frequent, alongside indirect content questions which again vary owing to the range of topics that are questionable under this function. Similarly, for Spanish, indirect content questions and indirect alternative questions are noteworthy. Again the range of forms is likely owing to the reflexivity of this function for asking a wide range of questions about ideas and concepts outside of the paper in which it is used.

One final area of consideration regarding question type and form centres on the discussion of the IFIDs used to raise questions; a detailed list of these is presented in Section 5.5. Like Blagojević and Misic-Ilic (2012), this study finds that verbs and nouns can serve as effective IFIDs for raising indirect questions. Overall, IFIDs used to raise indirect questions largely take interrogative verb forms such as “ask”, *interroger*, and *preguntar* in each language. The use of interrogative verbs is unsurprising given that these questions use these interrogative IFIDs in lieu of a question mark. As such, they appear to use the most straightforward substitution possible to raise indirect questions. The frequent use of “examine”, “explore” and *investigar* and *precisar* reflects Karttunen (1977) categorisation of verbs of acquiring knowledge. This shows English to be more concerned with getting information and adding something new, which does not reflect the formal-oriented nature designation the English language generally receives. However, this is consistent for the Spanish knowledge centred approach (Clyne 1994). For French, the use of *savoir* is of note as French seems less concerned with acquiring knowledge but retaining it, which is implicit in the verb which means “to know” (Karttunen 1977). For nouns, the use of “question” in each language is noteworthy in its use to raise indirect questions in each language. The importance of not constraining indirect questions by form cannot be understated. A limitation of corpus linguistics is that it is form driven. Therefore, to find IFIDs, limiting parts of speech can be problematic as this will limit the identification of potential language that can raise questions. Taking the ‘concealed question’ (Romero 2005) approach and not studying only interrogatives and wh-complements allowed this study to capture many more indirect question IFIDs. The process for identifying IFIDs in KIAP-EEFS is presented in Section 4.6.3 and is evaluated, methodologically, in Section 7.3.2. Overall, there is very little research on

IFIDs used to generate indirect questions in each language. Therefore, the IFIDs identified in this study make a valuable contribution to the description of IFIDs for indirect questions in each language. Further typological analyses of such IFIDs to identify the types of verbs and nouns used to raise questions would be interesting, as in Karttunen (1977). In terms of individual function, the spread of verbs and nouns makes it difficult to identify any patterns of use that generate strong relationships between functions and IFIDs. Although, using the identified IFIDs to study a larger corpus could generate interesting findings.

Overall, this section discussed the role of question types and forms in KIAP-EEFS. Each function of questions shows a different relation to question type and form in each language, meaning that form and function can be constrained. Recognising the perceived formality in academic writing in each language for their relative discourse community, this variation in question type and form use reflects the register of academic writing. Following Beun (2000), differing question use in different languages and for different purposes can be conceptualised as linguistic features of these academic genres. Overall, this section offers not only a detailed description of question forms in academic writing in English, French, and Spanish, but also insights into their semanto-syntactic constraints *i.e.* the relationship between function and form. The next section discusses the location of questions used in this study in terms of relevant literature.

7.2.5 Location of questions in the research article in English, French, and Spanish

Most questions in KIAP-EEFS occur within the thematic middle sections, followed by introduction, title, and conclusion. However, this finding is skewed owing to the greater amount of text and by proxy the greater opportunity to ask questions in the middle sections. Of note, proportionally, French and Spanish use more questions in introduction sections than English. However, such a finding is rather unrevealing and does little to highlight the behaviour of questions in academic writing. The one finding of note is the use of questions occurring more often in introductions in French and Spanish. This reflects Breeze (2012) who argues that academic genres in Spanish tend not to be so fixed and more open to variation. As such, in Spanish, questions may be less confined to particular spaces within the article. Similarly, for French, the discursive reflexivity the article affords may explain this greater variation within the text (Fløttum *et al.* 2006;

Bennet and Muresan 2016). However, when questions are considered in terms of their function, much more interesting findings are revealed on the relationship between function and location. Of the functions studied, getting attention and focussing the reader is constrained by titles and therefore cannot have location as a finding and organising the text exhibited the least variation with nearly all examples occurring within the middle sections. Taking the remaining findings from Chapter 6, the location of functions is discussed.

For framing the discourse, questions occur in the introduction and middle sections in English and in the introduction sections in French and Spanish. They are unusually frequent in introduction sections in each language. When compared to others like Curry and Chambers (2017), questions that frame the discourse in linguistics and economics present slightly different perspectives for English and French. In economics, English uses these questions in introduction sections and middle sections, while in linguistics they are largely confined to introduction sections. Conversely for French, these questions are largely confined to introduction sections in economics and, while in linguistics they are used in introduction and middle sections. This indicates that within the genre of the economics research article in English, it is not unusual to raise research questions outside of the introduction. This is not the case for French and Spanish. Overall, the presence of these questions in introductions in English is unsurprising given the evidence of role of question raising in introduction sections of economics research articles (Ryvityté 2003). This same question raising evidently carries across languages and the presence of such questions in introduction sections is characteristic of French (Curry and Chambers 2017) and Spanish (Lafuente-Millán 2014), despite Pérez-Llantada (2010) arguing that Spanish employs less participant-oriented metadiscourse in introductions. Similarly, for questions that create a niche a similar situation is found. These questions occur mainly in introduction sections in all three languages which again reflects the tendency to use questions that signal contribution early in research articles in English, French, and Spanish (Ryvityté 2003; Lafuente-Millán 2014; Curry and Chambers 2017).

English places questions that express attitudes and counter-claim in introduction and middle sections, while they occur mainly in middle sections in French, and middle and conclusion sections in Spanish. Overall, introduction sections in English and French make

unusually frequent use of this type of question while the same is true of conclusion sections in Spanish. The presence of so many varied question functions so frequently in introduction sections may be due to the larger space available in economics (Lakić 2010; 2015) to engage with the reader. This finding indicates that in English, if there is any volition to critique and rhetorically position others' work, it is likely to happen earlier in the article. For Spanish, it appears this criticism is enacted in conclusion sections. This is particularly interesting as conclusion sections are not consistently a part of the research article genre in Spanish. In Spanish, only 78% of the texts that compose the subcorpus contain conclusion sections. This indicates that Spanish tends to end research articles evaluating others' work and contributions. However this is only the case when conclusion sections are included in the article. For French, this fits somewhere in between.

Next, for setting up claims and protecting the writer, each language uses these questions in the middle sections. For English and French, these are unusually frequent in middle and conclusion sections, while Spanish sees their unusually frequent presence in conclusion sections. The evidence of these questions in conclusion sections is of interest as Dahl (2009) finds that claims are often signalled in conclusion sections of research articles. This same finding is reflected in French and Spanish. As such, claims have a distinctive relationship with conclusion sections more than any other. Finally, for asking real questions, English uses these questions in conclusion sections, French in middle and conclusion sections, and Spanish in introduction and middle sections. For French and Spanish, these questions are unusually frequent in conclusion sections. Interestingly, Hyland argues that such questions can close articles dramatically so as to maintain the readers' interest (2001) and while Salager-Meyer *et al.* (2003) argue that French and Spanish academic writing is more dramatic, in terms of real questions, English behaves more dramatically. Studies have also shown that conclusions are rich in metadiscourse in English economics (Dahl 2009; Tustin 2010b), which is also reflected in this question function. Overall, what these findings show is the degree of flexibility in each language for signalling gaps in knowledge and real questions. English tends to wait until the end to do this, while Spanish uses such questions much earlier in the text. French asks more of these type of question in the middle sections.

7.2.6 Passivity in academic writing in English, French, and Spanish

This section discusses the role of passivity in questions by drawing on relevant research on academic writing in English, French, and Spanish. Overall, there is evidence of very little passivity in any language. That is, while of course there is evidence of passivity in each language, most questions show active voice and clear agency. The lack of passivity in English academic writing is surprising given its established propensity to use passives (Tarone *et al.* 1981; Siewierska 1986; Shaw 1992; Lakić 1997; Martínez 2001; Martín 2003; Guerra and Guerra 2008; Baratta 2009; Bondi 2010; Flowerdew 2012; Amnuai 2019). This study finds the negligible use of passives in English is similar to Van Bonn and Swales (2007); for French and Spanish, passivity is infrequent while traditionally both languages have been found to use many passives in academic writing (Revilla 1980; Liddicoat 1992; 1993; 1995; Espinoza 1997; Martín 2003; Oliver del Olmo 2005; Fløttum *et al.* 2006; Pérez-Llantada 2010; Collins 2011; Salazar *et al.* 2013; Pérez and Ramiro 2015). In terms of function, there is evidence of stronger relationships between passivity and individual functions in each language. Questions that get attention and focus the reader and questions that organise the text showed no noteworthy use of passives; however the remaining five functions are worth discussing here.

For questions that frame the discourse, passivity is not frequent in general. However, Spanish exhibits an unusually frequent use of passivity for this function. As this function is rhetorically linked to posing research questions and largely occurs in introduction sections in Spanish, the use of passivity may be linked to the rhetorical use of passives to introduce topics, identified by Pérez-Llantada (2010). Moreover, Spanish is seen to be more impersonal in nature (Pérez-Llantada 2010) and passives have been found to be an effective way of demonstrating this impersonality (Salazar *et al.* 2013; Pérez and Ramiro 2015). For creating a niche, English makes unusually frequent use of passives. This function uses questions to establish a niche and these questions occur mostly in introduction sections. This use of passivity reflects Lakić's (1997) finding that in introduction sections in economics academic writing there is a propensity to make use of passivity in order to focus on the research itself and not the contribution of the writers to it. As such, the use of passives in this function allows writers to focus on the gap being identified and filled and not their role in identifying and filling it. Next, for expressing an attitude and counter-claiming, Spanish makes use of passivity, although it should be noted

that this function is so infrequent, it cannot be claimed that passivity is a feature of Spanish questions that express an attitude. Nonetheless, for expressing criticism, Spanish appears to evoke impersonality. For setting up claims, English, French, and Spanish all make use of passivity. For English, this use of passivity to set up claims is explicable where Bondi (2010), for example, finds that economics research articles use passive voices to introduce claims. Further studies show the value of passives is linked to the need to tone down the claims through minimising agency (Salager-Meyer 2011a). For French, the use of passives reflects Salager-Meyer's (2011a) view of French as being more impersonal and for Spanish, again, this use of passivity is consistent with the general literature on impersonality in academic writing (Salazar *et al.* 2013; Pérez and Ramiro 2015). However, the toning down of claims may be perceived as unusual in French and Spanish (Salager-Meyer 2011a). Equally this could be understood as a privileging of knowledge over agency as the focus of the sentence which is characteristic of French and Spanish (Bennett and Muresan 2016). Finally, for real questions, passivity is unusually frequent in English and French. As with many of the previously discussed examples, this reflects use of passives in both languages to tone down gaps in knowledge, and not link these gaps to the writer. Instead the focus is on the gap itself, and not the lack of knowledge on the writers' part.

Overall, passivity appears infrequent in each language with KIAP-EEFS; however looking at functions, there are interesting differences of note. Each language makes use of passivity in a different way. The findings demonstrate that there are clear relationships between function and passivity. Other studies on passivity have linked them to different functions (Lakić 1997), locations (Martín 2003), and considered them in terms of various forms (Espinoza 1997). This study, taking question function as the focus took a broad view of passivity and did not constrain by form, or passive function. In terms of location, this study was constrained to the corpus construction metadata and as such does not relate passivity to individual functions. While further elaboration of passivity in this way could be interesting in a larger data set, the data in this study were too few. As such, it was not considered worthwhile to further breakdown this use of passivity. Nonetheless, the findings presented and discussed show interesting variation across all three languages in terms of passivity. Next, in a similar process, tense use and verbal modality in questions are discussed, reflecting on the wider literature.

7.2.7 Tense and modality in academic writing in English, French, and Spanish

For tenses and verbal modality, present tense was most used in all languages. This corroborates long standing research such as Lackstrom, Selinker, and Trimble (1972) and Malcolm (1987) who find present simple tenses to be used for general claims. For questions that set up claims, there is a clear overlap. Similarly, others like Oster (1981) find present simple used in the creation of positive relationships between existing research and proposed research findings. Beyond claims, present tense can allow for the expression of neutrality, which also explains the overall propensity in English to use the present tense in questions. In terms of economics, Lakić (1997) finds that 55% of verbs take present simple tenses and in Deny *et al.* (2017), 73.6% of all verbs are in the present tense. The present is used to present facts (Lakić 1997) and also can refer to the past (Biber *et al.* 1999; Swales and Feak, 2012). There is a clear reflexivity in the present tense in English, which explains its frequent use in questions.

For French, present tense is also used most frequently, again owing to its reflexivity. This reflexivity refers to the use of present tense to cite research (Benichoux *et al.* 1985), present is used to refer to general information (Liddicoat 1992; 1993; 1995), and can also refer to the past (Burrough-Boenisch 2003). The range of question functions makes use of this reflexive nature of present tenses in French, which explains the general convergence towards the use of present tenses in French.

For Spanish this focus on present tenses again is also long standing (Espinoza 1997). Present tenses are most common in Spanish (Perez and Ramiro 2015) and are often used to describe processes (Lorés-Sanz 2011a; 2011b). However, despite the overall convergence in each language towards present tense use, each function did demonstrate more varied uses of tenses across languages, other than getting attention and focussing the reader, which almost exclusively used present tense. Functions are now discussed, focusing on languages and functions that used tenses other than present tense.

For framing the discourse, Spanish makes unusually frequent use of verbal modality. This function centres on the raising of questions that guide research, and as modality has been found to create impersonality and demonstrate stance (Pérez-Llantada 2010; Lorés-Sanz 2011a; 2011b), the role of modality in this function can serve to ensure the focus is on the content and not the researchers. This corresponds to Spanish as a content-oriented culture

(Cuenca 2003; Oliver del Olmo 2004). Next, for organising the text, English makes unusually frequent use of past tenses, French makes unusually frequent use of future tenses, and Spanish makes unusually frequent use of past and future forms. For English, the use of questions in organising the text serve to help readers navigate complex texts. As such, the past tense can be used, following Oster (1981), to build on previous research and engage with literature. The past is used to refer to specific research, and as such, questions that organise the text can focus on specific areas of research by exploiting tense. Moreover, as these questions are used to create connections between ideas, the use of past tenses creates supportive relationships between ideas, following Shaw (1992). Overall, from this finding, it can be inferred that questions that organise the text often refer to the past and this is especially noteworthy given the overall lower frequency of past tenses in English (Biber *et al.* 1999; Winkler and Metherell 2012). For French the unusually frequent use of future tenses indicates that connections made in texts see location in the text as a proxy for time in some way. The questions signal what will be discussed in the following sections, for example. There is little research on the role and function of future tenses in French academic writing as it is not very much used; however, its relevance for this function is notable. Finally, for Spanish, this function also sees future and past forms as unusually frequent. Following Martín (2003), the past tense is common in methodology sections and although it is not clear from this corpus what the exact relationship between methodology and questions is, they do occur in the middle section where this function is most common. Lorés-Sanz (2011a) also finds that past and future tenses are used in methodology sections of business management research articles. This may explain the occurrence in KIAP-EEFS in Spanish. Overall, there is a greater range of tenses evident in Spanish, while English, and more specifically, French, do not exhibit such range.

The next function considered is creating a niche. English makes notable use of modality and future tenses and Spanish uses past tenses. For English the presence of modality is typically high (Hyland 2004d) yet for future tenses, there is little research on their role in academic language. Interestingly, Nathan (2013) finds that business case studies use modality to propose hypothetical future actions and consequences. Therefore, for this function which highlights gaps to be filled, the focus is on the niche being filled. For Spanish, the focus appears to be on how the proposition fits in with what has already been done in the field. This relates to Lorés-Sanz (2011b) who finds that the past is used for

past acts and the perfect for a process in the past that relates to present and future for plans. Therefore, for creating a niche, Spanish appears to cross temporal boundaries in a way that English does not.

For expressing an attitude and counter-claiming, French and English use mainly verbal modality and make salient use of past tenses. The use of modality in both French and English is unsurprising, given the evaluative nature of the function. Moreover, modality is very frequent in English and French for hedging and evaluating claims, though more frequent in English than French (Salager-Meyer 2011a; Vold 2006). Modality indicates that this function allows for divergence from objective language in English (Parkinson and Adendorff, 2005) and seeks to express disagreement (Carrió-Pastor 2014) which is characteristic of this function of questions. This function is also rhetorically linked to the discussion of literature produced by others, which often necessitates modality (Deny *et al.* 2017). For French modality, this function reflects the use of evaluative modality (Becker and Remberger 2010) as modals accompany the judgement of others work. Therefore, it is unsurprising that modality is used in this function. For past tenses in English, like organising the text, this function of questions reflects a focus on specific research (Oster 1981) *i.e.* the research being evaluated, and the creation of relationships between different ideas. This has made use of past tenses in the literature review sections of papers quite common (Shaw 1992). For French, the use of past tenses raises the question of the location of these questions as past tenses like the imperfect are infrequent in French academic writing and are usually confined to methods sections (Liddicoat 1997). Owing to limitations of the data, the exact location of these questions is not known, but their use in methods is worth further investigation, perhaps in a larger data set.

For setting up claims, English uses mainly the past tense. However, it also exhibits salient use of future tense questions. French uses mainly past tenses and modals, and Spanish uses mainly modals, as well as a salient use of future tenses. Evidently, this function shows the most reflexivity in tense use. The past tense allows writers to raise claims in relation to specific research (Oster 1981) and create relationships between different ideas that form the basis of the claim (Shaw 1992). In this way, it is reasonable to argue that English uses past tenses to create doubt surrounding research and to propose solutions to that doubt. For future tenses, there is little existing research to draw on. Given the nature

of this function, claims using the future appear to address problems that will impact the future in some way. For French, claims' uses of modality signal doubt expressed through the use of modal verbs (Becker and Remberger 2010; Salager-Meyer *et al.* 2003). Therefore, creating claims through modality helps achieve this function in French. The past tense use in French is relatively infrequent and unusual, as past tenses have largely been described as constrained to methodology sections (Liddicoat 1997). For Spanish, the modality in this function is noteworthy as it reflects its more authoritative nature (Salager-Meyer *et al.* 2003).

Finally, real questions in French and English occur with modal verbs and English makes unusually frequent use of future tenses. Overall, for futures, there is little research to consider. The convergence of future and modality relates to Nathan's (2013) work in which modality is used to propose hypothetical future actions and consequences. Modality use in this function behaves somewhat like expressing an attitude. Overall, modality is very frequent in English and French academic writing (Salager-Meyer 2011b; Vold 2006) and in this function appears to be used to hedge (Hyland 2004d) and express doubt (Salager-Meyer 2003; 2011b) in order to signal a gap in knowledge on the writers' part.

Overall, as well as the many interesting tense-related differences that occur across languages in terms of function, what emerges as noteworthy is the similarity in the frequent use of present tense. Other relevant findings surrounding less studied tenses in academic writing in each language also pose valuable contributions to the current state of the art. The next section reviews the findings from the final criterial feature of this analysis: sentence type.

7.2.8 Question sentence type in academic writing in English, French, and Spanish

Simple and complex question sentence structures are most common in KIAP-EEFS. For English, complex questions are more common than simple questions and this is a similar finding to Li and Ge (2009) who find that complex sentences are also a feature of academic writing in medicine; however, others like and Chen (2019) find more compound sentences overall in doctoral theses which may indicate generic differences in sentence structure between research articles and other academic genres. The use of subordination is also reflected in Deveci and Nunn (2018) and overall, the evidence of simple sentences

also corresponds to Martín (2003) and Bennett (2009) who validate Dunleavy's (2003) advice for simple and straightforward grammar in English academic writing. However, it must be noted that in the case of questions, complex sentence structures are most common. This may be due to the presence of indirect questions which usually necessitate subordination. Owing to the constative approach taken in this study, inter-equivalence variation has not been studied. Therefore, this explanation remains anecdotal for now, though further tests with sentence types as *tertium comparationis* could reveal such interesting findings.

Like English, French uses more complex questions than simple which does not reflect Benichoux *et al.*'s (1985) advice for simplicity and readability in writing. Instead, this prevalence of complex sentences could be due to what Bennett and Muresan (2016) refer to as the linguistic artistry of academic writing in French which tends to favour more complicated sentences (Clyne 1994). When compared to English, French appears to have a greater penchant for subordination (Van Bon and Swales 2007); however in this study of question sentence types, the reverse is true. Finally, for Spanish, simple question types are more common than complex question. This does not reflect Pérez-Llantada (2010) who argue that complex sentences and subordination are particularly characteristic of academic Spanish in the context of business disciplines. Overall, questions in Spanish largely use simple clauses, which does not correspond to Bennett and Muresan's (2016, p.96) identification of Spanish as, 'elaborate', 'ornate', 'repetitious', and filled with 'flowery' sentences and subordination. This study, like Bennett and Muresan (2016), finds that verbless sentences are a feature of academic writing in French and Spanish and to a lesser degree, English. However it appears to be a small feature. Of course, all of the findings in this study are understood as representative of the language of questions. Therefore, this study cannot account for overall structure of sentences in KIAP-EEFS.

Overall, there is very little research on the sentence type and structure in academic writing in English, French, and Spanish, especially in the case of economics academic writing. When considering sentence type in questions and differences across functions, there is no research of note. Here the relationship between sentence type and question function is discussed to identify any particular linguistic features of questions in academic writing.

Starting with getting attention and focussing the reader, English uses mainly simple sentences, which are unusually common. This reflects Martín's (2003) and Bennett's (2009) designation of English as containing simple sentence structures. Given this question's limitation to existing within titles, it may also be that the simplicity is owing to journal limitations. For French and Spanish, complex questions feature highly which corresponds to Pérez-Llantada (2010) and Van Bon and Swales (2007). Therefore, questions to get attention and focus the reader differ across each language in terms of sentence type. For framing the discourse, of note is English's use of complex and compound-complex sentences which do not reflect the findings in the literature on sentences in English academic writing (Martín 2003). However, this may be owing to the many indirect questions found to be used in this function. For French and Spanish, compound sentences are noteworthy, where coordinating clauses build on information. This reflects the combination of independent aims nested within research questions in French and Spanish. For organising the text, English makes notable use of compound sentence types, creating relationships between independent lines of query. However, French shows complex-compound sentence types are used to organise the text which reflects Bennett and Muresan (2016) evaluation of French academic writing as overly complicated. For creating a niche, compound-complex sentences are unusually common in French and Spanish. English makes unusually frequent use of compound sentence types. Overall this indicates that questions used to create a niche and identify gaps in research tend to use more complicated sentences with multiple clauses. Given their focus on identifying gaps and proposing how to fulfil them, it is likely that this function necessitates more space to conduct its purpose. Next for expressing an attitude, English and Spanish use compound-complex questions which reflect the multiple moves typically involved in this evaluative function. For setting up claims, English makes use of compound-complex questions and for real questions, French makes use of compound-complex questions. This again shows important differences in the linguistic description of questions and the space required for the effective construction of questions that perform these functions.

Overall, there is little existing research on sentence types in academic writing in each language. That being said, there is a general convergence towards complex and simple questions with English, while French uses more complex questions and Spanish uses

more simple questions. Beyond this, each function exhibits a degree of variation across languages.

7.2.9 Contribution to research on questions in academic writing in English, French, Spanish and contrastive linguistics

This section discussed the findings of this PhD in terms of the wider literature on academic writing and makes clear the contributions this research makes to scholarship on questions as devices of reader engagement in academic writing in English, French, and Spanish. New information has been offered on the use of questions in economics research articles in each language, as well as their function. Moreover, details on their length, type and form, location of use, use of passivity, use of tense and modality, and their sentence type has been discussed. This discussion has referred to relevant research that positions the findings of this thesis in the wider context of academic writing. Next, consideration is given to the contributions this research makes to literature on the multilingual academy and corpus-based contrastive linguistics.

7.3 Contribution to research on the multilingual academy and corpus-based contrastive linguistics

Section 7.2 addressed the minutiae of this study in terms of findings on questions, their functions, and their length, type, form, location, voice, tense, and clausal construction in economics academic writing in English, French, and Spanish. This section takes these findings and considers their wider implications within research on the multilingual academy, in Section 7.3.1, and within research on corpus-based contrastive linguistics, in Section 7.3.2. The purpose of this section is to add to wider discourses in the field that have shaped this thesis and address gaps in the literature surrounding the lack of research on languages other than English and the contrastive study of more than two languages.

7.3.1 The multilingual academy

Within the wider discourses on academic writing, there are a number of themes, covered in Chapter 2, that are worth reconsidering. This refers to research on the impact of English on academic writing globally, as well as the role of culture in academic writing.

In Section 2.2.1, the role of English as the lingua franca of academia is discussed with the primary view to establish academic language in English, French, and Spanish as

comparable. However, that section raised a number of points surrounding the current face of academic language in French and Spanish and the challenges French and Spanish face under the threat of the ‘lingua frankensteinia’ (Phillipson 2008) or ‘tyrannosaurus rex’ (Swales 1997) that is English academic discourse. Much of that discussion centred on the recognition of the implicit or explicit privileging of the ways of thinking, constructing, and sharing knowledge in English-language academic culture (Bennett 2015), owing to the English language’s ‘privileged position’ within the academy (Fløttum *et al.* 2006, p.18). However, this thesis has identified a number of ways in which the corpus data can account for an alternative perspective to the imminent domain loss and epistemicide predicted in the literature (Bennett 2015).

While the English language has become a rule setter for other languages in the academy (Connor 1996), this does not account for the agency of academic writers in other languages nor for their ability to flout and alter established rules. In this thesis, the contrastive analysis of questions has identified a number of similarities and differences in academic writing in English, French, and Spanish. This study used contrastive approaches to ensure comparability of language, community, and genre, and determined that the research article, which is inherently linked to English language empiricist approaches (Bennett 2015), is a shared genre across languages. While the research article genre may originate from the English language academy, there is evidence in this study and the wider literature to show that both French and Spanish writers exhibit cultural and academic traditions in their writing in research articles.

As discussed in Section 7.2.1, questions as reader engagement markers occur to varying degrees across English, French, and Spanish. They occur across a range of texts and their differences in frequency reflect disciplinary and cultural facets of academic writing. For example, like Fløttum *et al.* (2006) and Grossmann and Tutin (2010), this study finds evidence of less metatext in French when compared to English. Furthermore, like Valero-Garcés (1996) there is evidence that English contains more metatext than Spanish. In terms of question function, Section 7.2.2 saw questions as reader engagement markers occur to varying degrees across English, French, and Spanish. For example, French uses more questions to express an attitude and counter-claim than English, and English uses more than Spanish. This function is rhetorically linked to expressing criticism (Hyland

2002a) and therefore, following Salager-Meyer *et al.* (2003) and Salager-Meyer and Ariza (2004), its presence is unsurprising as French is typically considered more critical. In terms of passivity in Spanish, questions that frame the discourse reflect Spanish language cultural practices. For example, earlier, in Section 7.2.6, it was argued that the use of passivity in Spanish is related to demonstrating impersonality, which is a common practice in Spanish academic writing (Salazar *et al.* 2013; Pérez and Ramiro 2015). Overall, throughout this chapter, differences across languages have been explained in terms of differences in academic cultural traditions. Therefore, what these examples show is that academic writing in the research article in French and Spanish is culturally different to that of English. Grouping the evidence of persisting academic traditions with the lack of a restrictive publish or perish culture in French and Spanish academia (Pérez-Llantada *et al.* 2011) and the well-established body of knowledge published in French and Spanish peer-reviewed journals (Breeze 2015), it is arguable that Bennett's 'monoculture of thought' (2015, p.11) is not truly imminent. Recalling MacKenzie (2015), this finding reflects arguments in the literature that writers' should be allowed to reflect a range of styles of writing in internationally published literature. Across languages, this seems, at least, to be the case.

Of course, it would be negligent to ignore the evidence of French and Spanish convergence towards English academic language norms (Ciapuscio and Otañi 2002; Salager-Meyer *et al.* 2003) and domain loss in academic publishing in French and Spanish (Ferguson 2007). Publication houses and academic journals as gatekeepers account for much of this problem and even within the specific context of economics (Henshall 2018). In writing this thesis, it became clear that research that is published in French and Spanish is often difficult to access (Carter-Thomas and Jacques 2017). Moreover, with fewer and fewer studies published in French and Spanish every year (Ferguson 2007), and conducted on French and Spanish academic language, there is a real risk of creating an academic diglossia (Ferguson 2007). Therefore, more research is needed on academic language in French and Spanish in order to document their academic language, acknowledge the complexities and ways of thinking inherent therein, and add to and develop the sparse collection of studies on academic discourse in French and Spanish. Publication houses and academic journals should also reconsider their publishing policy and further account for the internationalisation of research, as Henshall

(2018) identifies, in order to allow for more ways of presenting research and in more languages.

The final point to be discussed in this section refers to cultural identity in academic writing. In Section 2.2, the role of cultural identity in academic language was discussed and although there is a dearth of this research, it is rather consistent in viewing English as reader-oriented and formal-oriented, and French and Spanish as writer-oriented and content-oriented (Pisanski-Peterlin, 2005; Ädel 2006; Hyland and Salager-Meyer 2008; Amiryousefi and Rasekh 2010). However, such a view is not entirely supported by the findings of this study. Although questions as reader engagement are more common in English than French and Spanish, and more common in French than Spanish, each language demonstrated evidence of questions as reader engagement. Moreover, throughout the discussion of French and Spanish in this chapter, there is ample evidence to suggest that they both can behave as writer-responsible and formal-oriented languages *e.g.* the very frequent use of questions to organise the text in French and Spanish. While Ciapuscio and Otañi (2002) and Salager-Meyer *et al.* (2003) document a slow convergence towards English language norms in French and Spanish academic writing, which may explain this use of reader-oriented discourse, it seems an oversimplification to designate a language as being fixed as either writer- or reader-responsible. More reasonably, it seems that English, French, and Spanish are each writer-responsible and formal oriented, albeit to varying degrees. It would be interesting to consider whether different audiences would necessitate different degrees of reader orientation or what other variables can impact the decisions writers make to engage readers. Based on this study, to identify French and Spanish as not reader-friendly would be considered by Spack (1997) as culturally insensitive. A more specific and balanced claim would be that, in each language, writers consider their readers when writing academic texts. Based on this study, in descending frequency, English, French, and Spanish academic writers in economics research articles appear to attempt to engage readers through the use of rhetorical questioning.

7.3.2 Corpus-based contrastive linguistics

With regard to corpus-based contrastive linguistics, Section 3.6 identified some key gaps in the literature. The first pertained to the importance of relativism in the comparison of

languages and the subsequent need to conduct contrastive studies of at least three languages. The second gap relates to the lack of attention generally paid to contrastive rhetoric theories of comparability in corpus-based contrastive studies. This section briefly returns to these issues.

Given the well espoused view of contrastive analysis as the study and contrasting of two languages (Aijmer and Lewis 2017), this study has addressed the call from the literature to study three or more languages (Van der Auwera 2012). The argument made by Van der Auwera is that by including a third language it is possible to ‘make a point about the other two’ (2012, p.85). The results of this study very much support that view. As mentioned in Section 3.6, by comparing only two languages, this study would have had to present findings in a binary fashion. For example, this study finds that English contains more metadiscourse than French. Similarly, this study also finds that English contains more metadiscourse than Spanish. Initially, by only comparing English with French and English with Spanish, reader engagement findings are presented in a binary. These binary views support the identification of languages as reader- or writer-responsible for example, concepts which are problematised for being too polarising in Section 7.3.1.

A binary view supports these designations as without further languages as points of comparison, it is impossible to see degrees of reader engagement. However, with three languages, English appears to contain more metadiscourse than French, while French contains more metadiscourse than Spanish. Therefore, the extra point a third language allows to be made is one about relativism in contrastive studies which indicates that all contrastive studies are directional as they compare languages and cultures relative to criteria chosen as *tertia comparationis*. The advantage of multiple languages is well evidenced in Fløttum *et al.*’s (2006) work on the KIAP corpus and Carter-Thomas and Jacques’ (2017) work on the EIIDA project. Moving forward in the field, the results of this thesis recognise that contrastive analyses of three or more languages allow for more nuanced results. As such, where possible it is advisable that contrastive analyses incorporate more than two languages.

The second gap was concerned with the role of contrastive rhetoric theories in corpus-based contrastive analyses. The adequacy of the results of a contrastive study is directly related to comparability and the *tertium comparationis*. The underpinning values of the

tertium comparationis correspond to theories of representativeness in corpus linguistics, as they appear to be share core values surrounding the accurate and detailed definition of data samples. Therefore, it is surprising that so few corpus-based contrastive analyses in the literature consider concepts like the *tertium comparationis* or equivalence in study design. In fact, aside from Connor and Moreno (2005), Rabadán (2005; 2011), Fløttum *et al.* (2006). Rabadán and Izquierdo (2013), Carter-Thomas and Rowley-Jolivet (2013) and Izquierdo (2014), contrastive rhetoric theories were not apparent in the corpus linguistics literature on academic discourse. This is important because, while representativeness in the construction of a comparable corpus requires the consideration of sampling and the capturing of detailed metadata across languages (McEnery and Xiao 2008), corpus linguistics is not necessarily concerned with comparing like with like at all times. In fact, many corpus studies are designed to compare very different data (Lindquist and Levin 2000). However, in order to identify corresponding language in contrastive studies, comparing like with like is essential, as there can only be one assumed correspondence. In this study of questions as reader engagement, it was necessary to rigorously interrogate the data to ensure comparability. For example, in Section 2.2, the problematising of the comparability of English academic language with that of French and Spanish was not derived from issues of representativeness. It was derived of issues of comparability. These issues arise as, if the languages are not comparable, it becomes impossible to determine correspondences. Therefore, specifying the discipline of economics, the focus discourse community, and the genre of the research article allowed for effective comparison across languages.

Not only were contrastive rhetoric theories important in the validation of the corpus data, but they were useful in ensuring a rigorous analysis. As the *tertium comparationis* pertained to the shared use of questions as reader engagement, and the shared function of questions as reader engagement, it was important that there were no further constraints. However, as this study of questions focused on the identification of direct and indirect questions in KIAP-EEFS, there was potential for methodological circularity. This is because corpus searches are based on searches of form and as question form is a variable and finding in this study, contrastive linguistic theory determined that it was necessary to identify a means of searching for ‘illocutionary force indicating devices’ (Flöck and Geluykens 2015, p.7) that would signal both direct and indirect questions. This involved moving

away from a traditional ‘vertical-reading methodology’ (Aijmer and Rühlemann 2014, p.5) towards a function-to-form approach in corpus linguistics – the details for the identification of direct and indirect questions in KIAP-EEFS are presented in Section 4.2.3. Vertical-reading is not sufficient as, when studying the nature of engagement markers, ‘it is often necessary to look at context to understand the functions when taking a function-to-form approach’ (Curry and Chambers 2017, p.10).

Other ways in which contrastive linguistic theory informed the corpus approach was in the identification of equivalences. In the election of equivalences, frameworks developed by contrastive linguistic theorists like Krzeszowski (1990) were valuable for this study. By linking linguistic features of research articles to Krzeszowski’s (1990) equivalences, this study could effectively identify how to test the *tertia comparationis* across English, French, and Spanish. As such, equivalence allowed for the identification of criteria that can determine correspondences across languages.

Overall, by incorporating and merging approaches on contrastive and corpus linguistics, this study was able to produce robust and dependable findings that can identify clear correspondences in question use across languages. Within the wider field of corpus-based contrastive linguistics, it would be advisable for more studies to consider the contrastive theories underpinning their studies.

7.4 Conclusion

This chapter has discussed the findings of this PhD. First, focussing on academic writing in each language, questions were discussed in terms of frequency, function, question length, type and form, location, use of passivity, use of tense and verbal modality, and question sentence type. Following that, the contributions to research on the multilingual academy and corpus-based contrastive analysis were discussed. The following Chapter, Chapter 8, presents a summary of the contributions of this PhD while also discussing the pedagogical implications it proposes.

Chapter 8

Implications and future directions

8 Implications and future directions

8.1 Introduction

This chapter serves to conclude this research by reflecting on the research questions and considering the implications of the findings of this thesis for language teaching and learning. First, in Section 8.2, the contributions of this study are outlined, reflecting on the research questions that have guided this study. As this thesis has delivered linguistic descriptions that enrich current understandings of the academic practices of academic writers in economics in English, French, and Spanish, it is worth considering the possible implications these findings have for teaching and learning academic writing. Therefore, Section 8.3 considers the pedagogical implications of this research in terms of teaching academic writing, English for academic purposes, *français langue académique* and *español con fines académicos*, and corpus applications to language teaching and learning. Finally, to conclude this thesis, Section 8.4 presents some worthy follow-up studies which will build on and add to this research.

8.2 Contributions: answering the research questions

This section answers the core research questions and other relevant questions that guided this research.

1. To what extent does the presence of questions as reader engagement correspond in English, French, and Spanish economics research articles?

Questions as reader engagement are a feature of economics academic writing in each language. There is clear evidence that English uses questions as reader engagement more than French, and French more than Spanish. Moreover, the difference in usage between English and Spanish is statistically equivalent. Their functional distribution, locations, uses of passivity, question type and IFID, and use of tense and verbal modality, for example, show great correspondence across languages. However, question form and question sentence type show some differences. Overall, the view of this research is that the *tertium comparationis* corresponds across languages, and there are several shared features. The differences that occur do not impact the use of questions as reader engagement devices. Rather, these differences impact the presentation of questions. Chapter 5 presents the findings of this thesis.

2. To what extent does question function as reader engagement correspond in English, French, and Spanish economics research articles?

For the most part, functions are similarly distributed within languages, and English tends to use more than French, and French more than Spanish. However, this is not always the case, as each function exhibits differences across languages. It is the view of this thesis that question functions largely correspond across languages; however, within individual functions the way they correspond formally or syntactically, for example, can vary. That being said, the differences that occur do not impact the function of questions as reader engagement, rather they enrich them.

As mentioned in Section 1.4.1, other relevant questions emerged that were necessary to answer, in order to situate the context of this research. While this study adds to current knowledge on the use of questions in academic writing in economics research papers in English, French, and Spanish, the findings of this research can also offer commentary on wider issues in academic discourse and corpus-based contrastive linguistics. The remainder of this section returns to these questions and briefly summaries the contributions made to wider issues in the field.

1. What is the relationship between academic writing in English, French, and Spanish and how does this relationship impact the comparability of academic writing across the three languages?

In Section 2.2, relevant themes and studies were discussed to answer this question. In summary, the role of English as a lingua franca of academia was discussed with a focus on how it can negatively impact academic language in French and Spanish. Furthermore, consideration was given to culture and tradition in academic writing in order to illustrate how both the global context and the cultural origins that shape academic language can make comparison across languages challenging. In Section 7.3.1, these topics were reconsidered in order to identify how the results of this study can respond to them by considering the evidence of academic cultural traditions in economics research articles in English, French, and Spanish, and the polarising nature of classifications like reader- and writer-responsible languages.

2. What constitutes the discipline of economics in this study? Is it the same discipline in each language or does economics refer to something different in English, French, and Spanish?

This question was addressed in Section 2.3. Through a discussion of the literature on discipline, economics was established as a discipline in its own right. Identifying discipline in terms of its social nature as an academic tribe allowed this study to position the data in KIAP-EEFS as pertaining to the economics disciplinary discourse community. Through this definition of discipline, it was possible to position economics as a discipline that spans language. The discussion of literature throughout Chapter 2 and the discussion of findings throughout Chapter 7 make clear the importance of discipline as a variable in academic writing research. Overall, this study has added to studies on economics in three languages.

3. What is the nature of the discourse community studied herein? Is it one discourse community spanning English, French, and Spanish or does this study analyse three separate discourse communities in English, French, and Spanish? What is the impact of these conceptualisations of discourse community on the comparability of academic writing across the three languages?

In establishing the economics discourse community in Section 2.3, it became clear that working across languages required the application of a focus discourse community which is not limited by language, time, or space. In Section 7.3.2, the importance of ensuring comparability across languages through the identification of a focus discourse community was reconsidered. This drew attention to the value that contrastive linguistics offers corpus linguistics in terms of linguistic theory. That is because contrastive linguistics requires deep interrogation of data to ensure like is compared with like, which strengthens the quality of the data. The conclusion of this thesis is that within each language there is discourse community while across the three languages there is one focus discourse community. Being able to join the community means the comparison did not rest on comparing different groups. Rather the focus was on comparing different members of the same group.

4. To what extent is the genre of the research article comparable across English, French, and Spanish both textually and contextually? How do theories of genre identify the research article? Is the economics research article a disciplinary genre and is it comparable in English, French, and Spanish?

In Section 2.4-2.10, genre and the research article in English, French, and Spanish were considered. These sections focused on establishing the genre as one that spans language. This was achieved through consideration of both text and context in order to identify that

the economics research article in English, French, and Spanish shares social and communicative goals, as well as textual features defined by register. While the discussion in Chapter 2 showed that there were many differences across languages, fundamentally, the research article genre in each language served similar purposes and did so by similar means. In Chapter 7, the discussion of the findings in Section 7.2 interrogated the textual and metadiscoursal features of each language. This showed how the research article in each language, despite exhibiting differences, shared many rhetorical and textual features across languages.

5. What theories of comparability can ensure faithful comparisons in contrastive linguistics? What are the roles of the *tertium comparationis* and equivalence in contrastive studies? How do these concepts inform the approach to contrastive analysis employed herein?
6. What value does corpus linguistics offer to contrastive analysis? What research exists that merges these two fields? How do the findings of this study's corpus-based contrastive analysis relate to the current state of the art?

In Section 3.3, theories underpinning contrastive linguistics are discussed. Of these theories, the *tertium comparationis* and equivalence were prominent for their role in ensuring quality in the data and research design. This was followed by a discussion of corpus linguistics in Section 3.5. For corpus linguistics, theories of representativeness were important, as were methodological processes. As discussed in Section 7.3.2, contrastive linguistic theories were integral to this study and they merged well with corpus linguistic approaches. With advances in corpus linguistics largely occurring within statistical and quantitative movements, this research placed more emphasis on interrogating the quality of data, an area in need of revitalisation in corpus linguistics. A recommendation of this study is for more corpus-based contrastive analyses to draw from theory in contrastive linguistics.

7. What differences and similarities occur in the presence of questions in terms of frequency, function, question length, question type and form, question location, use of passivity in questions, use of tense and verbal modality in questions, and question sentence types in English, French, and Spanish?

Chapter 5 answers this question in detail. Based on the *tertium comparationis* of a shared use of questions as reader engagement, questions were analysed based on the above mentioned equivalences. Given that the *tertium comparationis* of shared question use was proven through the statistical analysis in Section 5.2, how questions correspond became

of more interest. Section 5.8 shows that there are many ways in which questions correspond to one another and these correspondences are composed of both similarities and differences. Based on these similarities and differences, Section 5.9 shows a potential correspondence. The results of these findings are discussed throughout Chapter 7.

8. What differences and similarities occur in question functions in terms of frequency, question length, question type and form, question location, use of passivity in questions, use of tense and verbal modality in questions, and question sentence types in English, French, and Spanish?

Chapter 6 answers this question in detail. Based on the *tertium comparationis* of shared question function as reader engagement, questions were analysed based on the above mentioned equivalences. Given that the *tertium comparationis* of shared question function was proven through the almost identical proportional distribution of question functions across languages, how question functions correspond became of more interest. Sections 6.2-6.8 show that each function behaves differently within and across languages. In the conclusion of each section on function, both frequent and salient features are presented. These are the maximally similar ways in which question functions correspond to one another across languages. Based on the similarities and differences established, Chapter 7 discussed the results of these findings.

8.3 Pedagogical implications

This section considers the value of comparable corpora, like KIAP-EEFS, for informing language teaching and the teaching of academic writing in English, French, and Spanish. As such, this section considers the findings of this thesis and their possible contributions to language learning in terms of research on genre, interpersonal writing, sentence level studies, as well as corpora for directly and indirectly informing teaching materials. While there are those who do not believe first language to be an important variable in English language publishing (Hyland 2016), at the very least the influence of first language on academic writing has largely been found to be non-negligible (Carter-Thomas and Rowley-Jolivet 2013, p.112). Therefore, the identification of differences across languages, such as those presented in this thesis, may prove useful in informing the teaching of language for academic purposes.

One area to which research on academic writing in English, French, and Spanish contributes is writing pedagogies in genre studies. Research, for example by Kuteeva and

Negretti (2016), shows that interdisciplinary analyses of student genre perceptions can offer valuable insights to teaching. Similarly, Hyland's (2015a) work on professional writing in genres and discipline demonstrates the importance of genre descriptions for language teaching. Biber and Conrad (2009, p.124), for example, propose a general register in academic prose that acts as a core across disciplines, which offers value for language teaching. Other studies on register are more applicable to discipline specific research such as Zhonghua's (2015) discipline specific study of register in terms of the applicability of register theory in teaching English for specific purposes. Moreover, from this perspective, work such as Hyland's (2004a; 2004c) study of discipline specific discourse communities furthers this view. Similarly in French, research has identified the need to focus on discipline, genre and register, as well as the differences therein (Cavalla and Grossmann 2005; Jisa 2004; Rinck 2011) as key to informing *français langue académique* and the development of discourse competence. Furthermore, for *español con fines académicos*, Hyland and Salager-Meyer (2008, p.320) identify that in languages like Spanish, the composition of a genre and register can differ substantially to others such as English. Moreover, Suarez Tejerina (2005) considers how text structure can influence the register and language used within a text. Building on these studies, this research has identified generic, disciplinary, and register specific features of question use in English, French, and Spanish. From a genre perspective, the different use of question functions in different locations within the research article indicates generic differences. These are different in economics than in linguistics, and in terms of register, all linguistic features that differ across languages are of note *e.g.* the different use of question forms. This indicates that while each language holds a formal academic register, writers of these registers present that formality differently.

The register of the genres can also be governed by the differing use of interpersonality in academic writing which are important in academic writing pedagogies. Like Lee and Deakin (2016, p.32), the results of this study can equip academic writers with sufficient and pertinent linguistic resources in order for them to situate themselves and their readers in their work. At the genre level, language pedagogy presents learners with 'superficial or abstract guidelines or formulas, scattered sentence examples, or a focus on vocabulary out of its discursive context' (Chang and Schleppegrell 2011, p.140). However, this alone is insufficient. Academic writing pedagogy, through investigation via appraisal

framework (Martin and White 2005), can move to equip students with the awareness of their authorial stance (Põldvere *et al.* 2016) and their readers' positioning (Arias and Román 2016). Similarly, O'Sullivan (2010) studies the acquisition of academic French and the ways in which stance taking can be better achieved. Lafuente-Millán (2014) looks at reader engagement and cultural contexts in business management articles in English and Spanish and presents useful insights into differences in interpersonal communication in academic writing in English and Spanish. As in the case of English for academic purpose and *français langue académique, español con fines académicos* has seen increased attention paid to the importance of teaching interpersonal writing, as writers are seen to be more impersonal and less interpersonal (Asención-Delaney 2014, p.240). Overall, these linguistic items can be problematic for learners where explicit reference to linguistic items that realise interpersonal communication are not made apparent at the genre level (Hood 2006, p.38), this is especially true in contexts where first language norms differ from target language norms. This is, of course, level dependent, as the teaching of academic writing at advanced levels would particularly require a shift to more complex and specific items (Chang and Schleppegrell 2016; Liu and Li 2016). Like these many studies, this research has delivered a detailed description of the role of questions as interactional metadiscourse and reader engagement devices in academic writing. As such, the findings regarding their frequency in each language, and functional distribution within each language can directly inform language pedagogies. It is understandable that the function of interpersonal communication within specific genres becomes important to academic writers and writing pedagogy; however, deeper still, the way in which interpersonality is realised textually is not always obvious, explicit, or directly translatable (Hood 2006, p.38). Further analyses of lexico-grammar become important in classifying, grouping, and sorting these markers into chunks, grammatical constructs or lexical bundles for example.

Focussing on the textual features studied in this thesis, there are contributions to be made to academic writing pedagogies. Research on lexical bundles for example has identified that academic writing can profit from chunks of vocabulary, such as 'the nature of the' or 'our goal is to' which can usefully be taught as set items (Biber and Barbieri 2007, pp.279-280). For informing *français langue académique*, research has centred on collocations (Cavalla 2008) both general *i.e. remettre en question* and specific *i.e. émettre une*

hypothèse or, following Tutin (2013), lexical bundles as discourse markers *i.e. dans un premier temps*. Similarly, for *español con fines académicos*, research has discussed the value of discourse markers to Spanish academic writing *i.e. es bien sabido que* or self-mentions *i.e. describimos*. (Escudero and Swales 2011, pp.61-63) while others like Cendón (2008) have studied lexical and verbal collocation in medical academic Spanish with expressions such as *realiza en* and *realizar por* which can boast useful pedagogical implications. Again, like these many studies, this research can make similar contributions. For example, the new findings surrounding question length, type and form, use of passivity, use of tense and verbal modality, and question sentence type offer detailed linguistic descriptions of questions that could be applied to the teaching of academic writing. Moreover, the detail on IFIDs used to raise indirect questions give very clear descriptions of verbs and nouns regularly used to raise questions which could have direct classroom applications.

All of these contributions to language pedagogy can be applied to English for academic purposes, *français langue académique*, and *español con fines académicos* contexts and can borrow from broader WAC, WID, and ACLIT approaches, see Section 1.3.1 for more on these school of teaching of academic writing. The specific value of the findings is to instil students with tacit knowledge about academic writing (McLeod and Miraglia 2001; Kruse 2013) and build genre awareness (Devitt 2004) most relevant to university contexts. (Harper and Vered 2016; Tuck 2016). Beyond these pedagogical applications, there are also corpus-informed pedagogical applications worth considering.

At a basic level, attested examples of natural language use in the context of academic writing in English, French, and Spanish can prove useful in indirectly informing language pedagogy. For example, following Hunston (2002), McEnery and Xiao (2005), and Römer (2011), the findings presented in this thesis surrounding IFIDs and grammatical structures, which published researchers use to pose questions in their research, can inform specialised grammars and dictionary entries for academic writing. Moreover, their inclusion in coursebooks and syllabi English for academic purposes, *français langue académique*, and *español con fines académicos* contexts (Mishan 2005; Timmis 2015) allows learners to use materials that ‘present vocabulary, grammar and functions students encounter most often in [academic] life’ (McEnery and Xiao 2011 p.367) and participate

in exercises built to develop discourse competencies for engaging readers in their writing. Such corpus-informed course books already exist *e.g.* *Touchstone* (McCarthy *et al.* 2004-2006), or, for an English for academic purposes example, *Unlock 2nd ed.* (Adams *et al.* 2019). Both of these course book series use corpus findings to inform various elements of their content.

For language testing, the role of corpora in language testing and assessment could be considered a relatively emerging field (Callies and Götz 2015, p.1) and corpora are now being used ‘to make inferences about language ability and capacity for’ learners to use a language (Chapelle and Plakans 2013. p.241). This is ushering a movement towards a data-driven approach to language assessment (Callies and Götz 2015, p.2). As such, the findings of this research could be valuable for informing academic skills tests to evaluate reader engagement in academic writing *i.e.* indicating what should be expected and what should not in terms of questions in academic writing.

For teachers of academic writing, these findings could be valuable for informing their professional development. This view reflects evidence that incorporating corpus linguistics into teacher training has been seen to boast extremely positive results, where ‘corpus evidence [can] address teachers’ questions’ (Tsui 2006, p.57) and raise awareness. Researchers like O’Keeffe and Farr argued that corpora can be used as tools to improve teachers’ knowledge, efficacy, and insight and, in so doing, develop a teaching expertise (2012, p.342). Knowledge, for example, may be acquired through increased awareness of how corpora work. Researchers such as Frankenberg-Garcia (2012) highlight how task-based approaches to teaching teachers about corpus linguistics can raise awareness for example, and such tasks could focus on questions in academic writing, based on this study. This improved awareness could help improve teaching efficacy as seen in studies such as Abdelkader *et al.*’s (2010) study which highlights how corpora can help improve teaching efficiency in online Arabic learning environments. On a contrastive level, incorporating corpus-based contrastive analysis into teacher training, where possible, could encourage and allow multilingual language teachers to exploit their own language learning as a tool of language analysis and teaching. This could raise language awareness and awareness of learner needs. Therefore, replicating elements of

this research could be of interest to teachers wishing to use corpora as part of their professional development.

Finally, for direct applications of corpus linguistics, data-driven learning approaches could be informed by this research (Johns 1991; 1994). Data-driven learning or ‘the learner as researcher’ (Bernardini 2006, p.16) takes an inductive approach to language learning (Gavioli 2005, p.28), putting learners in contact with much more language than traditional resources could ever hope to do (Chambers 2010, p.345). In data-driven learning, learners first observe the language to be studied, then they classify what they observed, focusing on the most important information. Finally, through generalisation, the learner can identify rules for the item being studied through grammatical analysis (Hadley 2002) or lexical and lexico-grammatical analysis (Timmis 2015, p.133). In this way, learners act as researchers (Hunston 2002, p.170) and are often guided by a series of questions or instructions and facilitated by the teacher whose role in such tasks is less directive (Gavioli 2005, p.30). Research has shown that interest in this type of learner led activity is increasing (Granath 2009, p.60), potentially owing to its efficacy for learning being substantiated by Boulton and Cobb’s (2017) seminal meta-analysis. To exemplify the application of this research to data-driven learning, Lee and Swales (2006) is a good model. They report on a project in which novice writers (non-native speaker postgraduate researchers) were asked to create a small corpus of articles in their discipline and analyse them with the aim to improve their writing. This produced generally positive results. Such a corpus could also be examined by the novice researchers to study how the published authors express aspects of stance and engagement, such as questions, in their writing. Alternatively, a teacher of academic writing could use such a corpus as a source of examples to illustrate the aspects of academic writing covered in the course. These applications could employ the use of a monolingual specialised corpus of expert writing or a comparable corpus of expert and/or novice writing, as in the case of Lee and Swales (2006).

In the case of specialised and multilingual corpora, approaches in data-driven learning can similarly be applied (Granger 2003, p.24; Almazova and Kogan 2014, p.248). For specialised language, the materials are simply more defined to serve the specific needs of the learners. Researchers such as Almazova and Kogan (2014) have discussed the

incorporation of data-driven learning into English for specific purpose courses. Similarly, Marzá (2014) looks at data-driven learning in teaching English for tourism. Beyond specialised corpora, corpus-based contrastive analyses of multilingual corpora boast similar extensive uses. They are seen to be particularly consciousness-raising (Granger 2003, p.24) and what is more, they can incorporate a particular approach to data-driven learning known as reciprocal learning. Using multilingual corpora in reciprocal learning, language learners take charge of their learning and the teacher becomes more of a material provider (Hunston 2002, p.184). Learners, in a multilingual classroom can use these corpora to teach each other as ‘each is a native speaker of the other’s target language’ (Boulton 2011, p.174), and exploit parallel corpora deductively to test preconceptions or answer questions, or inductively to explore language and learn more incidentally (Schmied 2011, p.168). Such approaches could be applied to the data used in this study, for example, by building an activity on IFIDs. In such an activity, learners could investigate how to raise indirect questions in their own disciplines and contexts.

Overall, all of these direct and indirect applications show how a small, specialised corpus can inform the writing of novice researchers in myriad ways. However, Kubota and Lehner (2004) caution the uncritical use of contrastive linguistics to inform language pedagogy. This study has shown the uneven nature of languages across domains and the ostensive connection between academic language and identity. Therefore, language pedagogies, based on contrastive research, should not diminish the value of the source language and should address head on the roles of rhetoric, culture, power, and discourse in writing. That being said, the use of bilingual and multilingual comparable corpora in language learning and teaching has received relatively little attention outside the context of translation studies (Laviosa 2002). Aston (1999, p.314) called for the application of comparable corpora for informing language pedagogy. Ten years later Granger (2009, p.9) continued to make this call, noting that the application of multilingual comparable corpora to language teaching warrants further investigation.

8.4 Looking forward

Moving forward, there are many opportunities for further study and exploitation of these data. While it is impossible to delineate all possibilities, a number of examples are

expressed here in order to identify the various directions and next steps in research in the fields to which this thesis responds.

- The data could be retested, changing the *tertium comparationis* to something like question type or form. This would give greater detail on questions from further perspectives.
- Other opportunities exist in modifying this study by adopting a divergent analysis and allowing for the inclusion of other metadiscourse markers in order to identify correspondences in each language.
- A further study could place a focus across disciplines which may help better distinguish between cultural traits and disciplinary discourse community traits of language use.
- Further research could focus on better identifying the relationship between different equivalences.
- The data could be modified. For example, the study could be repeated on a larger corpus; compared to reference corpora of other academic disciplines and genres; could be repeated on corpora of academic writing from different stages in the academic writing career such as BAWE (Nesi *et al.* 2008).
- Taking a larger corpus, more detailed statistical measures could be applied to identify significant relationships across the data.
- Given that this research has focused on writers' use of devices that engage readers, it would be worth studying readers' perceptions of these devices and the impact they have on their reading. This would prove invaluable in identifying the relationship between research on writing and research on reading. This extension of the study is clearly something a corpus study cannot afford, but would be well suited to research in psycholinguistics.
- Finally, using direct corpus applications, this study could be transformed to produce a data-driven learning guidance for language learners who wish to master the use of reader engagement.

Throughout the thesis, possible follow-up studies have been signalled where relevant and in this section, many of these have been reconsidered. At a more fundamental level, going forward, more work is needed on French and Spanish academic writing. Similarly, more

work is needed on better merging the fields of corpus and contrastive linguistics. It is the goal of this researcher to use the time following this PhD to further advance the field.

Reference List

Reference List

- Abdelkader, A., Boumiza, D.S. and Braham, R. (2010) 'An online Arabic learning environment based on IMS-QTI', in *Proceedings - 10th IEEE International Conference on Advanced Learning Technologies, ICALT 2010*, 116-118, available: <http://dx.doi.org/10.1109/ICALT.2010.40>.
- Adams, K., Baker, L., Brinks Lockwood, R., Dimond-Bayir, S., Jordan, N., Kennedy, A.S., Lansford, L., Lewis, M., O'Neill, R., Ostrowska, S., Pathare, E., Pathare, G., Peterson, S., Russell, K., Sowton, C., Westbrook, C., and White, J. (2019) *Unlock 2nd Edition* (6 levels), Cambridge: Cambridge University Press.
- Adamska-Sałaciak, A. (2010) 'Examining equivalence', *International Journal of Lexicography*, 23(4), 387-409, available: <http://dx.doi.org/10.1093/ijl/ecq024>.
- Adamska-Sałaciak, A. (2013) 'Equivalence, synonymy, and sameness of meaning in a bilingual dictionary', *International Journal of Lexicography*, 26(3), 329-345, available: <http://dx.doi.org/10.1093/ijl/ect016>.
- Ädel, A. (2006) *Metadiscourse in L1 and L2 English*, Amsterdam: John Benjamins.
- Agence bibliographique de l'enseignement supérieur (2016a) *Catalogue SUDOC*, available: <http://www.sudoc.abes.fr/?COOKIE=U10178,Klecteurweb,D2.1,Ec2b886e4-113,I250,B341720009+,SY,A%5C9008+1,,J,H2-26,,29,,34,,39,,44,,49-50,,53-78,,80-87,NLECTEUR+PSI,R109.125.17.154,FN> [accessed 09 Nov 2016].
- Agence bibliographique de l'enseignement supérieur (2016b) *Thèses.FR*, available: <http://www.theses.fr/en/> [accessed 09 Nov 2016].
- Aguirre, L. A. (2010) 'La construcción de la intersubjetividad en los informes de práctica profesional de Ingeniería', in Castel, V. M. and De Severino, L. C., eds., *La renovación de la palabra en el bicentenario de la Argentina. Los colores de la mirada lingüística*, Mendoza: Universidad Nacional de Cuyo. Facultad de Filosofía y Letras, 47-54.
- Aijmer, K. and Altenberg, B. (2013) *Advances in corpus-based contrastive linguistics: studies in honour of Stig Johansson*, Amsterdam: John Benjamins.
- Aijmer, K. and Lewis, D. (2017) 'Introduction', in Aijmer, K. and Lewis, D., eds., *Contrastive analysis of discourse-pragmatic aspects of linguistic genres (Vol. 5)*, Cham: Springer, 1-12..

- Aijmer, K. and Rühlemann, C. (2014) 'Introduction. Corpus pragmatics: laying the foundations', in Aijmer, K. and Rühlemann, C., eds., *Corpus pragmatics*, Cambridge: Cambridge University Press, 1-28.
- Al-Sibai, D.M. (2004) "'Not to be": the decline of contrastive analysis pedagogy', *Contrastive Analysis*, 523, 1-7.
- Alès, C., Arena, R., Brandt-Grau, A., Chaabane, N., Cortes, G., Crespin, R., Didier, E., Fretel, J., Gardey, D., Guermeur, I. and Gueye, L. (2016) *Rapport de prospective du Conseil scientifique de l'Institut des sciences humaines et sociales du CNRS*, Paris: Centre National de la Recherche Scientifique.
- Almazova, N. and Kogan, M. (2014) 'Computer assisted individual approach to acquiring foreign vocabulary of students major', *Lecture Notes in Computer Science*, Cham: Springer, 248-257.
- Altenberg, B. and Granger, S. (2002) 'Recent trends in cross-linguistic studies', in Altenberg, B. and Granger, S., eds., *Lexis in contrast*, Amsterdam: John Benjamins, 3-50.
- Amiryousefi, M. and Rasekh, A.E. (2010) 'Metadiscourse: definitions, issues and its implications for English teachers', *English Language Teaching*, 3(4), 159-167.
- Ammon, U. (2006) 'Language planning for international scientific communication: an overview of questions and potential solutions', *Current Issues in Language Planning* 7(1), 1-30.
- Amnuai, W. (2019) 'Analyses of rhetorical moves and linguistic realizations in accounting research article abstracts published in international and Thai-based journals', *SAGE Open*, 9(1), 1-9.
- Amossy, R. (2005) *L'argumentation dans le discours*, Paris: Armand Colin.
- Angermüller, J. (2011) 'From the many voices to the subject positions in anti-globalization discourse: enunciative pragmatics and the polyphonic organization of subjectivity', *Journal of Pragmatics*, 43(12), 2992-3000.
- Anthony, L. (2019) AntConc (Version 3.5.8) [Computer Software], Tokyo, Japan: Waseda University. Available from <https://www.laurenceanthony.net/software>.
- Araújo, A.D. (2010) 'Mapping genre research in Brazil: an exploratory study', in Bazerman, C., Krut, R., Lunsford, K., McLeod, S., Null, S., Rogers, P. and Stansell, A., eds., *Traditions of writing research*, New York: Routledge, 44-57.

- Arias, C.G. and Román, J.S. (2016) 'Conformación de ingroups y outgroups en el discurso de los lectores-comentaristas en blogs periodísticos', *Circulo de Linguistica Aplicada a la Comunicacion*, 65, 149-168.
- Asención-Delaney, Y. (2014) 'A multi-dimensional analysis of advanced written L2 Spanish', in Sardinha, T. B., Pinto, M. V. and Biber, D., eds., *Multi-dimensional analysis, 25 years on a tribute to Douglas Biber*, Amsterdam: John Benjamins, 239-269.
- Askerhave, I. and Swales, J. (2001) 'Genre identification and communicative purpose: a problem and a possible solution', *Applied Linguistics*, 22(2), 195-212.
- Aston, G. (1999) 'Corpora use and learning to translate', *Textus*, 12, 289-314.
- Aston, G. (2001) 'Learning with corpora: an overview', in Aston, G., ed., *Learning with corpora*, Bologna: CLUEB, 7-45.
- Bachschmidt P. (1999) 'Contruction de l'argumentation dans l'article de recherche en mécanique: différences entre discours du francophone et de l'anglophone', *ASp. La Revue du Geras*, 23, 197-208.
- Baker, P. (2012) 'Introduction', in Baker, P., ed., *Contemporary corpus linguistics*, London: Continuum.
- Bakhtin, M.M., Holquist, M., McGee, V. and Emerson, C. (1986) *Speech genres and other late essays*, Austin: University of Texas Press.
- Ball, R. (2009) 'Scholarly communication in transition: the use of question marks in the titles of scientific articles in medicine, life sciences and physics 1966–2005', *Scientometrics*, 79(3), 667–679.
- Baltar, M., Gastaldello, M.E.T. and Camelo, M.A. (2009) 'School radio: socio-discursive interaction tool in the school', *Educational Studies in Language and Literature*, 9(2), 49-70.
- Baratta, A.M. (2009) 'Revealing stance through passive voice', *Journal of Pragmatics*, 41(7), 1406-1421, available: <http://dx.doi.org/10.1016/j.pragma.2008.09.010>.
- Bawarshi, A.S. and Reiff, M.J. (2010) *Genre: An introduction to history, theory, research, and pedagogy*, West Lafayette: Parlor Press.
- Bayar, M. (2007) *To mean or not to mean*, Damascus: Kadmons Cultural Coundation Distribution.

- Beaufort, A. (1997) 'Operationalizing the concept of discourse community: a case study of one institutional site of composing', *Research in the Teaching of English*, 31(4), 486-495.
- Becher, T. (1989) *Academic tribes and territories: intellectual enquiry and the cultures of disciplines*, Milton Keynes: Open University Press.
- Becher, T. (1994) 'The significance of disciplinary differences', *Studies in Higher education*, 19(2), 151-61.
- Becker, M.G. and Remberger, E.M., eds. (2010) *Modality and mood in Romance: modal interpretation, mood selection, and mood alternation*, Berlin: De Gruyter.
- Benesch, S. (1996) 'Needs analysis and curriculum development in EAP. An example of a critical approach'. *TESOL Quarterly*, 30, 723-738.
- Benesch, S. (2001) *Critical English for academic purposes*, Mahwah: Erlbaum.
- Benveniste, E. (1974) *Problemes de linguistique generale. Vol 2*, Paris: Gallimard.
- Benichoux, R., Michel, J. and Pajaud, D. (1985) *Guide pratique de la communication scientifique*, Paris: G. Lachurie.
- Bennett, K. (2009) 'English academic style manuals: a survey', *Journal of English for Academic Purposes*, 8(1), 43-54.
- Bennett, K. (2015) 'Towards an epistemological monoculture: mechanisms of epistemicide in European research publication', in Plo, A.R. and Pérez-Llantada, C., eds., *Volume 2 English as a scientific and research language. Debates and discourses*. Berlin: De Gruyter, 9-36.
- Bennett, K. and Muresan, L.M. (2016) 'Rhetorical incompatibilities in academic writing: English versus the romance cultures', *Synergy*, 12, 95-119.
- Bernardini, S. (2006) 'Corpora in the classroom: an overview and some reflections on future developments', in Sinclair, J.M., ed., *How to use corpora in language teaching*, Amsterdam: John Benjamins, 15-38.
- Beun, R. J. (2000) 'Context and form: declarative or interrogative, that is the question', in Bunt H. and Black, W., eds., *Abduction, belief and context in dialogue: studies in computational pragmatics*, Amsterdam: John Benjamins, 311-326.
- Biber, D. (2006a) *University language: a corpus-based study of spoken and written registers*, Amsterdam: John Benjamins.

- Biber, D. (2006b) 'Stance in spoken and written university registers', *Journal of English for Academic Purposes*, 5(2), 97-116, available: <http://dx.doi.org/10.1016/j.jeap.2006.05.001>.
- Biber, D. (2012) 'Representativeness in corpus design', in Biber, D. and Reppen, R., eds., *Corpus linguistics vol. 4*, Los Angeles: Sage, 3-35.
- Biber, D. and Barbieri, F. (2007) 'Lexical bundles in university spoken and written registers', *English for Specific Purposes*, 26(3), 263-286, available: <http://dx.doi.org/http://dx.doi.org/10.1016/j.esp.2006.08.003>.
- Biber, D. and Conrad, S. (2001) 'Register variation: a corpus approach', in Schiffrin, D., Tannen, D. and Hamilton, H. E., eds., *The handbook of discourse analysis*, Oxford: Blackwell Publishers, 175-196.
- Biber, D. and Conrad, S. (2009) *Register, genre, and style*, Cambridge: Cambridge University Press.
- Biber, D., Johansson, S., Leech, G., Conrad, S. and E., F. (1999) *Longman grammar of spoken and written English*, Harlow: Longman.
- Biber, D. and Reppen, R. (2012) 'Preface', in Biber, D. and Reppen, R., eds., *Corpus linguistics Vol. 1*, London: Sage.
- Birch-Bécaas, S. (1996) 'French researchers publishing in English. An analysis of a corpus of first drafts', *ASp. La Revue du Geras*, 14, 75-88.
- Bizzell, P. (1992) *Academic discourse and critical consciousness*, Pittsburgh: University of Pittsburgh Press.
- Bhatia, V.K. (1993) *Analysing genre: language use in professional settings*, London: Longman.
- Bhatia, V.K. (2002a) 'A generic view of academic discourse', in Flowerdew, J., ed., *Academic discourse*, New York: Routledge, 21-39.
- Bhatia, V.K. (2004) *Worlds of written discourse: a genre-based view*, London: Continuum.
- Blagojević, Savka and Misic-Ilic, Biljana. (2012) 'Interrogatives in English and Serbian academic discourse – a contrastive pragmatic approach', *Brno Studies in English*, 38, 17-35.
- Bloom, L., Merkin, S. and Wootten, J. (1982) "'Wh"-questions: linguistic factors that contribute to the sequence of acquisition', *Child Development*, 53(4), 1084-1092.

- BNC (2007) BNC XML Edition, distributed by Bodleian Libraries, University of Oxford, on behalf of the BNC Consortium. Available: <http://www.natcorp.ox.ac.uk/>
- Bocanegra-Valle, A. (2014) ‘English is my default academic language’: voices from LSP scholars publishing in a multilingual journal, *Journal of English for Academic Purposes*, 13, 65-77, available: <http://dx.doi.org/10.1016/j.jeap.2013.10.010>.
- Bondi, M. (2010) ‘Arguing in economics and business discourse: phraseological tools in research articles’, *Bulletin Suisse de Linguistique Appliquée*, 2, 219-234.
- Borg, E. (2003) ‘Discourse community’, *ELT Journal*, 57(4), 398-400, available: <http://dx.doi.org/10.1093/elt/57.4.398>.
- Boulton, A. (2011) ‘Data-driven learning: the perpetual enigma’, in Gozdz-Roszkowski, S., ed., *Explorations across languages and corpora*, London: Peter Lang, 563-580.
- Boulton, A. and Cobb, T. (2017) ‘Corpus use in language learning: a meta-analysis’, *Language Learning*, 67(2), 348-393.
- Bradley, P.T. and Mackenzie, I. (2004) *Spanish: an essential grammar*, London: Routledge.
- Breeze, R. (2012) *Rethinking academic writing pedagogy for the European university*, Amsterdam: Rodopi.
- Breeze, R. (2015) ‘Citing outside the community? An investigation of the language of bibliography in top journals’, in Plo, A.R. and Pérez-Llantada, C., eds., *Volume 2 English as a scientific and research language. Debates and discourses*. Berlin: De Gruyter, 37-58.
- Bres, J. (2001) ‘Dialogique, dialogisme, dialogise (marquers de)’, in Cetrie, C., Sibot, P. and Verine, B., eds., *Termes et concepts pour l’analyse du discours. Une approche praxématique*, Paris: Champion, 86-89.
- Brezina, V. (2018) *Statistics in corpus linguistics: a practical guide* Cambridge: Cambridge University Press.
- Brezina, V., Timperley, M., & McEnery, T. (2018) #LancsBox v. 4.x [computer software]. Available: <http://corpora.lancs.ac.uk/lancsbox>.
- Briscoe, E., Carroll, J., and Watson, R. (2006) ‘The second release of the RASP system’, in Proceedings of the coling/acl 2006 interactive presentation sessions, Sydney: COLING/ACL, 77–80.

- Bronckart, J.P. (1997) *Activité langagière, textes et discours: pour un interactionisme socio-discursif*. Paris: Delachaux et Niestlé.
- Burrough-Boenisch, J. (2003) 'Examining present tense conventions in scientific writing in the light of reader reactions to three Dutch-authored discussions', *English for Specific Purposes*, 22(1), 5-24.
- Buter, R. and van Raan, A.F. (2011) 'Non-alphanumeric characters in titles of scientific publications: an analysis of their occurrence and correlation with citation impact', *Journal of Informetrics*, 5(4), 608-617.
- Callies, M. and Götz, S. (2015) 'Learner corpora in language testing and assessment: prospects and challenges', in Callies, M. and Götz, S., eds., *Learner corpora in language testing and assessment*, Amsterdam: John Benjamins, 1-12.
- Carrió-Pastor, M.L. (2014) 'Cross-cultural variation in the use of modal verbs in academic English', *Sky Journal of Linguistics*, 27, 23-45.
- Carter, R. and McCarthy, M. (2006) *Cambridge grammar of English: a comprehensive guide; spoken and written English grammar and usage*, Cambridge: Cambridge University Press.
- Carter-Thomas, S. and Chambers, A. (2012) 'From text to corpus: A contrastive analysis of first person pronouns in economics article introductions in English and French', in Boulton, A., Carter-Thomas, S. and Rowley-Jovilet, E., eds., *Corpus-informed research and learning in ESP: issues and applications*, Amsterdam: John Benjamins, 17-44.
- Carter-Thomas, S. and Jacques, M-P. (2017) 'Interdisciplinary and interlinguistic perspectives on academic discourse: the mode variable', *Chimera: Romance Corpora and Linguistic Studies*, 4(1), 1-11.
- Carter-Thomas, S. and Rowley-Jolivet, E. (2013) 'Citation from a cross-linguistic perspective: the case of French researchers publishing in English', *ESP across Cultures*, 10, 111-126.
- Cassany, D. (1999) 'Actitudes y valores sobre la composición escrita', *Alegria de Enseñar. La Revista para Maestros y Padres*, 40, 22-28.
- Cavalla, C. (2008) 'Les collocations dans les écrits universitaires: un lexique spécifique pour les apprenants étrangers', *Apprendre une langue de spécialité: enjeux culturels et linguistiques*, 93-104.

- Cavalla, C. and Grossmann, F. (2005) 'Caractéristiques sémantiques de quelques "Noms scientifiques" dans l'article de recherche en français', in Torodd, K., ed., *Akademisk prosa*, Bergen: KIAP Romansk institutt, 47-59.
- Cendón, B.M. (2008) 'Las colocaciones en la prosa académica médica: análisis contrastivo inglés-español', in Mellado Bianco, C., ed., *Colocaciones y fraseología en los diccionarios*, Frankfurt am Main: Peter Lang, 117-130.
- Cesteros, S.P. (2013) 'Enseñanza de español con fines profesionales y académicos y aprendizaje por contenidos en contexto universitario', in Gimeno Menéndez, F., Montoya Abad, B. and Mas i Miralles, A., eds., *"Studia linguistica in honorem" Francisco Gimeno Menéndez*, Alicante: San Vicent del Raspeig, 71-88.
- Chambers, A. (2010) 'What is data-driven learning?', in O'Keeffe, A. and McCarthy, M., eds., *The Routledge handbook of corpus linguistics*, London: Routledge Ltd, 345-358.
- Chambers, A. (2011) 'Language learning as discourse analysis: playing games in a corpus of French journalistic discourse', in Kübler, N., ed., *Corpora, language, teaching, and resources from theory to practice*, Bern: Peter Lang.
- Chang, P. and Schleppegrell, M. (2016) 'Explicit learning of authorial stance-taking by L2 doctoral students', *Journal of Writing Research*, 8(1), 49-79, available: <http://dx.doi.org/10.17239/jowr-2016.08.01.02>.
- Charteris-Black, J. and Ennis, T. (2001) 'A comparative study of metaphor in Spanish and English financial reporting', *English for Specific Purposes*, 20(3), 249-266, available: [http://dx.doi.org/10.1016/S0889-4906\(00\)00009-0](http://dx.doi.org/10.1016/S0889-4906(00)00009-0).
- Chau, M.H. (2012) 'Learner corpora and second language acquisition', in Hyland, K., Chau, M. H. and Handford, M., eds., *Corpus applications in applied linguistics*, London: Continuum, 191-207.
- Chen, Q. (2019) 'Theme-Rheme structure in Chinese doctoral students' research writing. From the first draft to the published paper', *Journal of English for Academic Purposes*, 37, 154-167.
- Cheng, W. (2012) *Exploring corpus linguistics: language in action*, New York: Routledge.
- Chesterman, A. (1998) *Contrastive functional analysis*, Amsterdam: John Benjamins.

- Chesterman, A. (2007a) 'Similarity analysis and the translation profile', in Vandeweghe, W., Vandepitte, S. and Van de Velde, M., eds., *The study of language and translation*, Amsterdam: John Benjamins, 53-66.
- Chesterman, A. (2007b) 'Similarity and difference in translation', in Arduini, S. and Hodgson, R., eds., *Similarity and difference in translation: proceedings of the international conference on similarity and translation*, Roma: Edizioni di storia e letteratura, 63-76.
- Chomsky, N. (1964) 'The development of grammar in child language: discussion', *Society for Research in Child Development*, 29(1), 35-42, available: <http://dx.doi.org/10.2307/1165753>.
- Chomsky, N. (2009) *Cartesian linguistics: a chapter in the history of rationalist thought*, Cambridge: Cambridge University Press.
- Ciapuscio, G. and Otañi, I. (2002) 'Las conclusiones de los artículos de investigación desde una perspectiva contrastiva', *Revista del Instituto de Investigaciones Lingüísticas y Literarias RILL*, 15, 117-133.
- Clyne, M. (1994) *Inter-cultural communication at work: cultural values in discourse*. Cambridge, Cambridge University Press.
- Cmejrkova, S. and Danes, F. (2011) 'Academic writing and cultural identity: the case of Czech academic writing', in Duszak, A., ed., *Culture and styles of academic discourse*, Berlin: De Gruyter, 41-62.
- Cobb, T. and Horst, M. (2005) 'Is there room for an academic word list in French', in Bogaards, P., ed., *Vocabulary in a second language: selection, aquisition, and testing*, Amsterdam: John Benjamins, 15-38.
- Coffin, C. and Donohue, J.P. (2012) 'Academic Literacies and systemic functional linguistics: how do they relate?', *Journal of English for Academic Purposes*, 11(1), 64-75, available: <http://dx.doi.org/10.1016/j.jeap.2011.11.004>.
- Collentine, J. (2014) 'Subjunctive in second language Spanish', in Geeslin, K.L., ed., *The handbook of Spanish second language acquisition*, Oxford: Wiley-Blackwell, 270-286.
- Collins (2011) *Easy Learning French Grammar*, New York: Harper Collins.
- Connor, U. (1996) *Contrastive rhetoric: cross-cultural aspects of second-language writing*, Cambridge: Cambridge University Press.

- Connor, U. (2002) 'New directions in contrastive rhetoric', *TESOL Quarterly*, 36(4), 493-510.
- Connor, U. (2003) 'Changing currents in contrastive rhetoric: implications for teaching and research', in Kroll, B., ed., *Exploring the Dynamics of Second Language Writing*, Cambridge: Cambridge University Press, 218-241.
- Connor, U. and Moreno, A.I. (2005) 'Tertium comparationis: a vital component in contrastive rhetoric', in Bruthiaux, P., Atkinson, D., Eggington, W., Grabe, W. and Ramanatan, V., eds., *Directions in applied linguistics: essays in honour of Robert B*, Bristol: Multilingual Matters, 153-164.
- Connor, U., Nagelhout, E. and Rozycki, W. (2008) *Contrastive rhetoric: reaching to intercultural rhetoric*, Amsterdam: John Benjamins.
- CSUC (2016) *TDX (Tesis Doctorals en Xarxa): TDX Home*, available: <http://www.tdx.cat/> [accessed 09 Nov 2016].
- Cook, J.M. and Plourde, D. (2016) 'Do scholars follow Betteridge's law? The use of questions in journal article titles', *Scientometrics*, 108(3), 1119-1128.
- Cop, M. and Hausman F.J. (1991) 'Collocations in the bilingual dictionary', in *An international encyclopaedia of lexicography*, Berlin: De Gruyter, 2775-2778.
- Cortes, V. (2008) 'A comparative analysis of lexical bundles in academic history writing in English and Spanish', *Corpora*, 3(1), 43-57, available: <http://dx.doi.org/10.3366/E1749503208000063>.
- Coxhead, A. (2000) 'A new academic word list', *TESOL Quarterly*, 34(2), 213-238.
- Cuenca, M.J. (2003) 'Two ways to reformulate: a contrastive analysis of reformulation markers', *Journal of Pragmatics*, 35, 1069-1093.
- Curry, N. and Chambers, A. (2017) 'Questions in English and French research articles in linguistics: a corpus-based contrastive analysis', *Corpus Pragmatics*, 1(3), 1-24, available: <http://dx.doi.org/10.1007/s41701-017-0012-0>.
- Dafouz, E. and Camacho-Miñano, M.M. (2016) 'Exploring the impact of English-medium instruction on university student academic achievement: the case of accounting', *English for Specific Purposes*, 44, 57-67.
- Dahl, T. (2004) 'Pragmatics of discourse. Textual metadiscourse in research articles: a marker of national culture or of academic discipline?', *Journal of Pragmatics*, 36(10), 1807-1825, available: <http://dx.doi.org/10.1016/j.pragma.2004.05.004>.

- Dahl, T. (2009) 'The linguistic representation of rhetorical function: a study of how economists present their knowledge claims', *Written Communication*, 26(4), 370-391.
- Darbellay, F. (2015) 'Rethinking inter- and transdisciplinarity: undisciplined knowledge and the emergence of a new thought style', *Futures*, 65, 163-174, available: <http://dx.doi.org/10.1016/j.futures.2014.10.009>.
- Davies, Mark (2010) 'The corpus of contemporary American English as the first reliable monitor corpus of English', *Literary and Linguistic Computing*, 25(4), 447-65.
- Day, R. and Sakaduski, N. (2011) *Scientific English: a guide for scientists and other professionals*, California: ABC-CLIO.
- Deignan, A. (2011) 'Searching for metaphorical patterns in corpora', in Baker, P., ed., *Contemporary corpus linguistics*, London: Continuum.
- Dendale, P. (2006) 'Three linguistic theories of polyphony/dialogism: an external point of view and comparison', *Sproglig polyfoni. Arbejdspapirer*, 5, 3-32.
- Deny A., Angkita K. and Almira F. (2017) 'The distribution of verb tenses and modals in journal articles' abstracts', *Southern African Linguistics and Applied Language Studies*, 35(3), 229-244.
- Dervin, F., Johansson, M. and Mutta, M. (2007) 'Ecriture académique: collaboration multimodale et "constructions identitaires" en FLA', *Synergies Europe*, 2, 93-107.
- Deveci, T. and Nunn, R. (2018) 'The use of relative clauses in humanities and social sciences research articles: a case study', *Linguistics and Literature Studies*, 6(1), 17-26.
- Devitt, A.J. (2004) *Writing genres*, Carbondale: Southern Illinois University Press.
- Dewey, M. (2007) 'English as a lingua franca and globalization: an interconnected perspective', *International Journal of Applied Linguistics*, 17(3), 332-354, available: <http://dx.doi.org/10.1111/j.1473-4192.2007.00177.x>.
- Dietz, G. (2001) 'The pragmatics of scientific titles formulated as questions', in: Kertész, A., ed., *Approaches to the pragmatics of scientific discourse*, Frankfurt Am Main: Peter Lang, 19-33.
- Donahue, C. (2002) 'The lycée-to-university progression in French students' development as writers', in Foster, D. and Russell, D.R., eds., *Writing and*

- learning in cross-national perspective. Transitions from secondary to higher education*, Urbana: NCTE, 134-191.
- Donahue, C. (2008) 'When is copying not copying. Plagiarism and French composition scholarship', in Eisner, C., Eisner, C. and Vicinus, M., eds., *Originality, imitation, and plagiarism : teaching writing in the digital age*, Ann Arbor: University of Michigan Press, 90-103.
- Du Plessis, H., Sehume, J. and Martin, L. (2014) *Concept and Application of Transdisciplinarity in Intellectual Discourse and Research*, Johannesburg: Real African Publishers.
- Ducrot, O. (1984) *Le dire et le dit*, Paris: Minuit.
- Dunleavy, P. (2003) *Authoring a PhD: how to plan, draft, write and finish a doctoral thesis or dissertation*, London: Palgrave Macmillan.
- Duszak, A. (1997) 'Cross-cultural academic communication: a discourse community view', in Duszak, A., ed., *Culture and styles of academic discourse*, Berlin: De Gruyter, 11-40.
- Eckart, T. and Quasthoff, U. (2013) 'Statistical corpus and language comparison on comparable corpora', in Sharoff, S., ed., *Building and Using Comparable Corpora*, Berlin: Springer, 151-166.
- Egan, T. (2014) 'Encoding thoroughness in English and French', in Alcaraz-Sintes, A. and Valera-Hernández, S., eds., *Diachrony and synchrony in English corpus linguistics*, Berlin: Peter Lang, 233-259.
- Escudero, M.P. and Swales, J.M. (2011) 'Tracing convergence and divergence in pairs of Spanish and English research article abstracts: the case of Ibérica', *Ibérica: Revista de la Asociación Europea de Lenguas para Fines Específicos*, 21, 49-70.
- Espinoza, A.M. (1997) 'Contrastive analysis of the Spanish and English passive voice in scientific prose', *English for Specific Purposes*, 16(3), 229-243.
- Eynsham, A. (c.955 - c.1010) *Grammatica latino-saxonica: una cumejusdem glossario Latino-Saxonica*, republished in Somner, W. (1659), Oxford: Oxonii.
- Evison, J. (2010) 'What are the basics of analysing a corpus', in O'Keeffe, A. and McCarthy, M., eds., *The Routledge Handbook of Corpus Linguistics*, London: Routledge, 122-135.

- Fairclough, N. (2003) *Analysing discourse: textual analysis for social research*, New York: Routledge.
- Farkas, D.F. and Bruce, K.B. (2010) 'On reacting to assertions and polar questions', *Journal of Semantics*, 27(1), 81-118.
- Farr, F. (2003) 'Engaged listenership in spoken academic discourse: the case of student-tutor meetings', *Journal of English for Academic Purposes*, 2(1), 67-85, available: [http://dx.doi.org/10.1016/S1475-1585\(02\)00035-8](http://dx.doi.org/10.1016/S1475-1585(02)00035-8).
- Farzannia, S. and Farnia, M. (2016) 'Metadiscourse markers in introduction sections of Persian and English mining engineering articles', *English for Specific Purposes World*, 49(17), 1-16.
- Ferguson, G. (2007) 'The global spread of English, scientific communication and ESP: questions of equity, access and domain loss', *Ibérica: Revista de la Asociación Europea de Lenguas para Fines Específicos*, 13, 7-38.
- Fisiak, J. (1980) *Theoretical issues in contrastive linguistics*, Amsterdam: John Benjamins.
- Flöck, I. and Geluykens, R.I.E. (2015) 'Speech acts in corpus pragmatics: a quantitative contrastive study of directives in spontaneous and elicited discourse', in Romero-Trillo, J., ed., *Yearbook of corpus linguistics and pragmatics 2015*, London: Springer, 7-37.
- Fløttum K. (2012) 'Variation of stance and voice across cultures', in Hyland K. and Guinda C.S., eds., *Stance and Voice in Written Academic Genres*, London: Palgrave Macmillan, 218-231.
- Fløttum, K., Dahl, T. and Kinn, T. (2006) *Academic voices across languages and disciplines*, Amsterdam: John Benjamins.
- Fløttum, K., Dahl, T., Didriksen, A.A. and Gjesdal, A.M. (2013) 'KIAP—reflections on a complex corpus', *Bergen Language and Linguistics Studies*, 3(1), 138-150.
- Flowerdew, J. (2002) 'Introduction: approaches to the analysis of academic discourse in English', in Flowerdew, J., ed., *Academic discourse*, Harlow: Longman, 1-18.
- Flowerdew, J. and Wang, S.H. (2015) 'Identity in academic discourse', *Annual Review of Applied Linguistics*, 35, 81-99.
- Flowerdew, J. and Forest, R.W. (2015) *Signalling nouns in English: a corpus-based discourse approach*, Cambridge: Cambridge University Press.

- Flowerdew, L. (2005) 'An integration of corpus-based and genre-based approaches to text analysis in EAP/ESP: countering criticisms against corpus-based methodologies', *English for Specific Purposes*, 24(3), 321-332, available: <http://dx.doi.org/http://dx.doi.org/10.1016/j.esp.2004.09.002>.
- Flowerdew, L. (2012) 'Grammar and the research article', in Chapelle, C.A., ed., *The Encyclopedia of Applied Linguistics*, Oxford: Wiley-Blackwell, 1-7, available: [10.1002/9781405198431.wbeal0473](http://dx.doi.org/10.1002/9781405198431.wbeal0473).
- Frankenburg-Garcia, A., Flowerdew, L. and Aston, G. (2013) 'Part III: corpora by language learners: learner language', in Frankenburg-Garcia, A., Flowerdew, L. and Aston, G., eds., *New trends in corpora and language learning*, London: Continuum, 181-262.
- Freedman, A. (2013) 'The what, when, where, why and how of classroom genres', in Petraglia, J., ed., *Reconceiving writing, rethinking writing instruction*, London: Routledge, 121-144.
- Fries, C.C. (1945) *Teaching and learning English as a foreign language*, Ann Arbor: University of Michigan Press.
- Friginal, E., Li, M. and Weigle, S.C. (2014) 'Revisiting multiple profiles of learner compositions: a comparison of highly rated NS and NNS essays', *Journal of Second Language Writing*, 23, 1-16, available: <http://dx.doi.org/10.1016/j.jslw.2013.10.001>.
- Fundación Dialnet (2016) *Portada de tesis - Dialnet*, available: <https://dialnet.unirioja.es/tesis> [accessed 09 Nov 2016].
- Gabrielatos, C. (2018) 'Keyness analysis: nature, metrics and techniques', in Taylor, C. and Marchi, A. eds., *Corpus approaches to discourse: a critical review* London: Routledge, 225-258.
- García Negroni, M.M. (2008) 'Subjetividad y discurso científico-académico: acerca de algunas manifestaciones de la subjetividad en el artículo de investigación en español', *Revista Signos*, 41(66), 9-31.
- Gardner, R. (2001) *When listeners talk*, Amsterdam: John Benjamins.
- Gardner, S. and Nesi, H. (2012) 'A classification of genre families in university student writing', *Applied Linguistics*, 34(1), 25-52.
- Garfinkel, H. (1967) *Studies in ethnomethodology*, Englewood Cliff: Prentice-Hall.

- Gavioli, L. (2005) *Exploring corpora for ESP learning*, Philadelphia: John Benjamins.
- Gea-Valor, M.L., Rey-Rocha, J. and Moreno, A.I. (2014) 'Publishing research in the international context: an analysis of Spanish scholars' academic writing needs in the social sciences' *English for Specific Purposes*, 36, 47-59.
- Ghivirigă, T (2012) 'Scientific articles in English in economics journals. A case study for Romania', *Procedia-Social and Behavioral Sciences*, 46, 4230-4235.
- Gilquin, G. (2008) 'Prototypical equivalence and lexical learning', Paper presented at the *Third International Conference of the German Cognitive Linguistics Association*, Leipzig 25-27 September.
- Gilquin, G. (2015) 'At the interface of contact linguistics and second language acquisition research: new Englishes and Learner Englishes compared', *English World-Wide*, 36(1), 91-124, available: <http://dx.doi.org/10.1075/eww.36.1.05gil>.
- Goddard, C. (1998) *Semantic analysis: a practical introduction*, Oxford: Oxford University Press.
- Goddard, C. (2012) 'Semantic primes, semantic molecules, semantic templates: key concepts in the NSM approach to lexical typology', *Linguistics*, 50(3), 711-743, available: <http://dx.doi.org/10.1515/ling-2012-0022>.
- Goddard, C. and Wierzbicka, A. (2008) 'Universal human contexts as a basis for contrastive linguistic semantics', in Gómez Gonzales, M., Mackenzie, J. and Gonzales Alvarez, E., eds., *Current trends in contrastive linguistics*, Amsterdam: John Benjamins, 205-226.
- Goffman, E. (1971) *Relations in public: microstudies of the public order*, New York: Basic Books.
- Gómez, I.F., Gómez, S.P., Silveira, J.C.P. and Garcia, J.F.C., (1998) 'Disciplinary variations in the writing of research articles in English', in Fortanet, I. and Dudley-Evans, T., eds., *Genre studies in English for academic purposes*, Valencia: Publicaciones de la Universitat Jaume, 59-78.
- Granath, S. (2009) 'Who benefits from learning how to use corpora?', in Aijmer, K., ed., *Corpora and language teaching*, Amsterdam: John Benjamins, 47-66.
- Granger, S. (2003) 'The corpus approach: a common way forward for contrastive linguistics and translation studies?', in Granger, S., Lerot, J. and Petch-Tyson, S.,

- eds., *Corpus-based approaches to contrastive linguistics and translation studies*, Amsterdam: Rodopi, 17-31.
- Granger, S. (2009) 'The contribution of learner corpora to second language acquisition and foreign language teaching', in Aijmer, K., ed., *Corpora and language teaching*, Amsterdam: John Benjamins, 13-32.
- Granger, S. (2010) 'Comparable and translation corpora in cross-linguistic research, design, analysis and applications', *Journal of Shanghai Jiaotong University*, 2, 14-21.
- Granger, S. (2012) 'How to use foreign and second language learner corpora', in Mackey, A. and Gass, S., eds., *Research methods in second language acquisition: a practical guide*, Malden: Blackwell, 7-29.
- Granger, S., Dagneaux, E. and Meunier, F. eds. (2002) *International corpus of learner English. Handbook*, Louvain: UCL Press.
- Gray, B. and Biber, D. (2012) 'Current conceptions of stance', in Hyland, K. and Sancho-Guinda, C., eds., *Stance and voice in written academic genres*, New York: Palgrave MacMillan, 15-33.
- Gries, S.T. (2010) 'Useful statistics for corpus linguistics', in Sanchez, A., ed., *A mosaic of corpus linguistics: Selected approaches*, Bern: Peter Lang, 269-291.
- Gries, S.T. (2012) 'Dispersion and adjusted frequencies in corpora', in Biber, D. and Reppen, R., eds., *Corpus linguistics vol. 4*, Los Angeles: Sage, 111-145.
- Gries, S.T. (2019) *Ten lectures on corpus linguistics with R: applications for usage-based and psycholinguistic research*, The Hague: Brill.
- Griffiths, C. and Parr, J.M. (2001) 'Language-learning strategies: theory and perception', *ELT Journal*, 55(3), 247-254.
- Grossmann, F. and Tutin, A. (2010) 'Evidential markers in French scientific writing: the case of the French verb *voir*', in Diewald, G. and Smirnova, E., eds., *Linguistic realization of evidentiality in European languages*, Berlin: De Gruyter, 279-308.
- Guan, S.H., Wang, Y.L. and Luo, L.J. (2006) 'An analysis on political speech from the perspective of appraisal system', *Journal of East China Jiaotong University*, 248, 98-101.

- Guerra, C.H. and Guerra, J.M.H. (2008) 'Discursive analysis and pragmatic metadiscourse in four sub-areas of economics research articles', *Ibérica, Revista de la Asociación Europea de Lenguas para Fines Específicos*, (16), 81-108.
- Haarman, H. and Holman, E. (2001) 'The impact of English as a language of science in Finland and its role for the transition to network society', in Ammon, U., ed., *The dominance of English as a language of science: effects on other languages and language communities*, Berlin: De Gruyter, 229-260.
- Hadley, G. (2002) 'Sensing the winds of change: an introduction to data-driven learning', *RELC Journal*, 33(2), 99-124.
- Haggan, M. (2004) 'Research paper titles in literature, linguistics and science: dimensions of attraction', *Journal of Pragmatics*, 36, 293-331.
- Halliday, M.A.K. (1978) *Language as social semiotic: the social interpretation of language and meaning*, Baltimore: University Park Press.
- Hamby, S. (2015) 'On scientific writing in the information era: tailoring papers for Internet searching and other 21st century realities', *Psychology of Violence*, 5(2), 103-111.
- Hamilton, C. and Carter-Thomas, S. (2017) 'Competing influences: the impact of mode and language on verb type and density in French and English scientific discourse', *CHIMERA: Romance Corpora and Linguistic Studies*, 4(1), 13-34.
- Harper, R. and Vered, K.O. (2016) 'Developing communication as a graduate outcome: using 'writing across the curriculum' as a whole-of-institution approach to curriculum and pedagogy', *Higher Education Research & Development*, 1-14, 688-701, available: <http://dx.doi.org/10.1080/07294360.2016.1238882>.
- Hartwell, L., Emmanuelle E-R. and Tutin, A. (2017) 'I think we need...: verbal expressions of opinion in conference presentations in English and in French.', *CHIMERA: Romance Corpora and Linguistic Studies*, 4(1), 35-60.
- Harwood, N. (2005) 'We do not seem to have a theory ... the theory I present here attempts to fill this gap': inclusive and exclusive pronouns in academic writing', *Applied Linguistics*, 26(3), 343-375, available: <http://dx.doi.org/10.1093/applin/ami012>.

- Hasan, R. (2009) 'The place of context in a systemic functional model', in Halliday, M.A.K. and Webster, J.J., eds., *Continuum companion to systemic functional linguistics*, London: Continuum, 166-189.
- Hausmann, F.J. (1989) 'Pour une histoire de la méta lexicographie', in *An international encyclopaedia of lexicography*, Berlin: De Gruyter, 216-224.
- Hausman F.J. and Werner R.O. (1991) 'spezifische bauteile und strukturen zweisprachigen wörterbücher: eine übersicht', in *An international encyclopaedia of lexicography*, Berlin: De Gruyter, 2729-2769.
- Hawkins, J.A. (1986) *A comparative typology of English and German: unifying the contrasts*, Austin.: University of Texas Press.
- Hawkins, R. and Franceschina, F. (2004) 'Explaining the acquisition and non-acquisition of determiner-noun gender concord in French and Spanish', in Prévost, P. and Paradis, J., eds., *The Acquisition of French in different contexts: focus on functional categories*, Amsterdam: John Benjamins, 175-206.
- Hawkins, R. and Towell, R. (2015) *French grammar and usage*, London: Routledge.
- He, M. and Abdul Rahim, H. (2019) 'Comparing engagement markers in economics research articles and opinion pieces: a corpus-based study', *Gema Online Journal of Language Studies*, 19(2), 1-14
- Heller, V. and Morek, M. (2015) 'Academic discourse as situated practice: an introduction', *Linguistics and Education*, 31, 174-186, available: <http://dx.doi.org/10.1016/j.linged.2014.01.008>.
- Hempel, S. and Degand, L. (2006) 'The use of sequencers in academic writing: a comparative study of French and English', *International Symposium: Discourse and Document*, 5(1), 35-39.
- Henshall, A.C. (2018) 'English language policies in scientific journals: signs of change in the field of economics', *Journal of English for Academic Purposes*, 36, 26-36.
- Hirschbühler, P. and Valois, D. (1992) 'Argument extraction out of indirect questions in French', in Goodluck H. and Rochemont M., eds., *Island constraints. Studies in theoretical psycholinguistics*, Dordrecht: Springer, 195-222.
- Higher Education Authority (2016) *HOME / Digital Repository Ireland*, available: <http://dri.ie/> [accessed 09 Nov 2016].

- Hinds, J. (1987) 'Reader versus writer responsibility: a new typology', in Connor, U. and Kaplan, R., eds., *Writing across languages: analysis of L2 text*, Reading: Addison Wesley, 141-152.
- Holmes, R. (1997) 'Genre analysis, and the social sciences: an investigation of the structure of research article discussion sections in three disciplines', *English for Specific Purposes*, 16(4), 321- 337.
- Holub, M., Kríž, V., Cinková, S. and Bick, E. (2012) 'Tailored feature extraction for lexical disambiguation of english verbs based on corpus pattern analysis', in *24th International Conference on Computational Linguistics - Proceedings of COLING 2012: Technical Papers*, Bombay: IIT Bombay, 1195-1210.
- Hood, S. (2006) 'The persuasive power of prosodies: radiating values in academic writing', *Journal of English for Academic Purposes*, 5(1), 37-49, available: <http://dx.doi.org/10.1016/j.jeap.2005.11.001>.
- Hood, S. (2010) *Appraising research: evaluation in academic writing*, New York: Palgrave Macmillan.
- Hosseini-Maasoum, S.M., Ghiasian, M., Roshan, B. and Shahidi, A.S. (2016) 'The transition process of semantic prosody to connotation in persian; a synchronic analysis', *Language Related Research*, 7(1), 17-38.
- Hu, G. and Cao, F. (2015) 'Disciplinary and paradigmatic influences on interactional metadiscourse in research articles', *English for Specific Purposes*, 39, 12-25.
- Huddleston, R. (1984) *Introduction to the grammar of English*, Cambridge: Cambridge University Press.
- Huise, J. (1624) *A perfect survey of the English tongue taken according to the use and analogie of the Latine*, London: Printed by Robert Young, for Iohn Rothwell in Pauls Church-yard at the signe of the Sunne, 1632.
- Hunston, S. (2002) *Corpora in applied linguistics*, Cambridge: Cambridge University Press.
- Hunston, S. and Thompson, G. (2000) *Evaluation in text: authorial stance and the construction of discourse*, Oxford: Oxford University Press.
- Hutchinson, T. and Waters, A. (1987) *English for Specific Purposes: a learning-centred approach*, Cambridge: Cambridge University Press.

- Hyland, K. (1998) 'Persuasion and context: the pragmatics of academic metadiscourse', *Journal of Pragmatics*, 30(4), 437-455.
- Hyland, K. (2001) 'Bringing in the reader: addressee features in academic articles', *Written Communication*, 18(4), 549-574, available: <http://dx.doi.org/10.1177/0741088301018004005>.
- Hyland, K. (2002a) 'What do they mean? Questions in academic writing', *Text*, 22(4), 529-557.
- Hyland, K. (2002b) 'Genre: language, context, and literacy', *Annual Review of Applied Linguistics*, 22, 113-135.
- Hyland, K. (2004a) *Disciplinary discourses: social interactions in academic writing*, Ann Arbor: University of Michigan Press.
- Hyland, K. (2004b) *Patterns of engagement: dialogic features and L2 undergraduate writing*, London: Continuum.
- Hyland, K. (2004c) 'Engagement and disciplinarity: the other side of evaluation', Del Lungo Camiciotti, G. and Tognini Bonelli, E., eds., *Academic discourse - new insights into evaluation*, Frankfurt Am Main: Peter Lang, 13-30.
- Hyland, K. (2004d) 'Disciplinary interactions: metadiscourse in L2 postgraduate writing', *Journal of Second Language Writing*, 13(2), 133-151, available: <http://dx.doi.org/http://dx.doi.org/10.1016/j.jslw.2004.02.001>.
- Hyland, K. (2005a) 'Stance and engagement: a model of interaction in academic discourse', *Discourse Studies*, 7(2), 173-192, available: <http://dx.doi.org/10.1177/1461445605050365>.
- Hyland, K. (2005b) *Metadiscourse: exploring interaction in writing*, London: Continuum.
- Hyland, K. (2008) 'Persuasion, interaction and the construction of knowledge: representing self and others in research writing', *International Journal of English Studies*, 8(2), 1-23.
- Hyland, K., (2009) *Academic discourse: English in a global context*, New York: Continuum.
- Hyland, K. (2010) 'Metadiscourse: Mapping interactions in academic writing', *Nordic Journal of English Studies*, 9(2), 125-143.

- Hyland, K. (2011) 'Academic discourse', in Hyland, K. and Paltridge, B., eds., *Continuum companion to discourse analysis*, New York: Continuum, 171-184.
- Hyland, K. (2015a) 'Genre, discipline and identity', *Journal of English for Academic Purposes*, 19, 32-43, available: <http://dx.doi.org/10.1016/j.jeap.2015.02.005>.
- Hyland, K. (2015b) *Teaching and researching writing: third edition*, Florence: Taylor and Francis.
- Hyland, K., (2016) 'Academic publishing and the myth of linguistic injustice', *Journal of Second Language Writing*, 31, 58-69.
- Hyland, K. and Hamp-Lyons, L. (2002) 'EAP: issues and directions', *Journal of English for Academic Purposes*, 1(1), 1-12, available: [http://dx.doi.org/10.1016/S1475-1585\(02\)00002-4](http://dx.doi.org/10.1016/S1475-1585(02)00002-4).
- Hyland, K. and Jiang, F.K. (2016) "'We must conclude that...': a diachronic study of academic engagement", *Journal of English for Academic Purposes*, 24, 29-42.
- Hyland, K. and Salager-Meyer, F. (2008) 'Scientific writing', *Annual Review of Information Science and Technology*, 42(1), 297-338, available: <http://dx.doi.org/10.1002/aris.2008.1440420114>.
- Hyland, K. and Tse, P. (2004) 'Metadiscourse in academic writing: a reappraisal', *Applied Linguistics*, 25(2), 156-177, available: <http://dx.doi.org/10.1093/applin/25.2.156>.
- Izquierdo, M. (2014) 'On Describing similarity and measuring equivalence in English-Spanish translation', *Meta: Journal des traducteursMeta:/Translators' Journal*, 59(1), 140-159.
- James, C. (1980) *Contrastive analysis*, Harlow: Longman.
- Jafari, S.S. (2014) 'A cross-linguistic study of english and persian prepositions', *Journal of Language Teaching and Research*, 5(3), 689-697, available: <http://dx.doi.org/10.4304/jltr.5.3.689-697>.
- Jiang, F. and Hyland, K. (2015) "'The fact that': stance nouns in disciplinary writing", *Discourse Studies*, 17(5), 529-550.
- Jisa, H. (2004) 'Growing into academic French', in Berman, R. A., ed., *Language development across childhood and adolescence*, Philadelphia: John Benjamins, 135-162.
- Johansson, S. (2007) *Seeing through multilingual corpora: on the use of corpora in contrastive studies*, Amsterdam; Philadelphia: J. Benjamins.

- Johns, T. (1991) 'Should you be persuaded: two examples of data-driven learning', in Johns, T. and King, P., eds., *Classroom concordancing elr journal*, Birmingham: University of Birmingham Press, 1-16.
- Johns, T. (1994) 'From printout to handout: grammar and vocabulary teaching in the context of data-driven learning', in Odin, T., ed., *Perspectives on pedagogical grammar*, Cambridge: Cambridge University Press, 293-313.
- JStor (2019) *Annales d'Économie et de Statistique*, available : <https://www.jstor.org/journal/annaeconstat> [last accessed 01 Nov 2019].
- Kalisz, R. (1981) *The pragmatics, semantics, and syntax of the English sentences with indicative trat complements and the Polish sentences with że complements: a contrastive study*, Gdańsk: Uniwersytet Gdański.
- Karpinska, L. (2015) *English-Latvian lexicographic tradition: a critical analysis*, Berlin: De Gruyter.
- Karttunen, L. (1977) 'Syntax and semantics of questions', *Linguistics and Philosophy*, 1(1), pp.3-44.
- Kanoksilapatham, B. (2012) 'Facilitating scholarly publication: genre characteristics of English research article introductions and methods', *The Southeast Asian Journal of English Language Studies*, 18(4), 5-19.
- Kanoksilapatham, B. (2015) 'Distinguishing textual features characterizing structural variation in research articles across three engineering sub-discipline corpora', *English for Specific Purposes*, 37, 74-86, available: <http://dx.doi.org/10.1016/j.esp.2014.06.008>.
- Kaplan, R. (1966) 'Cultural thought patterns in inter-cultural education', *Language Learning*, 16(2), 1-20.
- Kennedy, G.D. (2014) *An introduction to corpus linguistics*, London: Longman.
- Khedri, M., Ebrahimi, S.J. and Heng, C.S. (2013a) 'Interactional metadiscourse markers in academic research article result and discussion sections', *The Southeast Asian Journal of English Language Studies*, 19(1), 65-74.
- Khedri, M., Heng, C.S. and Ebrahimi, S.F. (2013b) 'An exploration of interactive metadiscourse markers in academic research article abstracts in two disciplines', *Discourse Studies*, 15(3), 319-331

- Khedri, M., Chan, S.H. and Helen, T. (2015) 'Interpersonal-driven features in research article abstracts: cross-disciplinary metadiscoursal perspective', *Pertanika Journal of Social Sciences & Humanities*, 23(2), 303-314.
- Khoutyz, I. (2015) 'Engagement in written academic discourse: a cross-cultural study of Russian and English research articles', *International Journal of Russian Studies*, 4(2), 135-160.
- Kilgariff, A., Baisa, V., Bušta, J., Jakubiček, M., Kovář, V., Michelfeit, J., Rychlý, P. and Suchomel, V. (2014) 'The Sketch Engine: ten years on', *Lexicography*, 1(1), pp.7-36.
- Klein, J.T. (2013) 'The transdisciplinary moment (um)', *Integral Review*, 9(2), 189-199.
- Klein, J.T. (2014) 'Interdisciplinarity and transdisciplinarity: keyword meanings for collaboration science and translational medicine', *Journal of Translational Medicine and Epidemiology*, 2(2), 1024-1031.
- Klein, P.D. and Boscolo, P. (2016) 'Trends in research on writing as a learning activity', *Journal of Writing Research*, 7(3), 311-350, available: <http://dx.doi.org/10.17239/jowr-2016.07.03.01>.
- Klein, W. (1986) *Second language acquisition*, Cambridge: Cambridge University Press.
- Kopczynski, A. (1984) 'Problems of quality in conference interpreting', in Fisiak, J., ed., *Contrastive linguistics: prospects and problems*, Berlin: De Gruyter, 283-300.
- Kordić, L. and Cigan, V. (2013) 'Teaching and learning foreign languages for legal purposes in Croatia', *Studies in Logic, Grammar and Rhetoric*, 34(1), 59-74.
- Kortmann, B. and Meyer, P.G. (1992) 'Is English grammar more explicit than German grammar after all?', in Mair, C. and Markus, M., eds., *New departures in contrastive linguistics*, Innsbruck: University of Innsbruck Press, 155-166.
- Koutsantoni, D. (2004) 'Attitude, certainty and allusions to common knowledge in scientific research articles', *Journal of English for Academic Purposes*, 3(2), 163-182.
- Koutsantoni, D. (2007) *Developing academic literacies: Understanding disciplinary communities' culture and rhetoric*, Oxford: Peter Lang.
- Krifka, M. (2014) 'Embedding illocutionary acts', in Roeper, T. and Speas, M., eds., *Recursion, complexity in cognition*, London: Springer, 125-155.

- Kruse, O. (2013) 'Perspectives on academic writing in european higher education: genres, practices, and competences', *REDU: Revista de Docencia Universitaria*, 11(1), 37-58.
- Krzeszowski, T.P. (1990) *Contrasting languages: the scope of contrastive linguistics*, Berlin: De Gruyter.
- Kubota, R. and Lehner, A. (2004) 'Toward critical contrastive rhetoric', *Journal of Second Language Writing*, 13(1), 7-27.
- Kuhn, T.S. (1970) *The structure of scientific revolutions*, Chicago: University of Chicago Press.
- Kuteeva, M. and Negretti, R. (2016) "'Graduate students' genre knowledge and perceived disciplinary practices: Creating a research space across disciplines', *English for Specific Purposes*, 41, 36-49, available: <http://dx.doi.org/10.1016/j.esp.2015.08.004>.
- Lackstrom, J. E., Selinker, L. and Trimble. L. P. G. (1972) 'Grammar and Technical English', *English Teaching Forum*, 13(4), 250-260.
- Lado, R. (1957) *Linguistics across cultures: applied linguistics for language teachers*, Ann Arbor: University of Michigan Press.
- Lafuente-Millán, E. (2014) 'Reader engagement across cultures, languages and contexts of publication in business research articles', *International Journal of Applied Linguistics*, 24(2), 201-223, available: <http://dx.doi.org/10.1111/ijal.12019>.
- Lakić, I. (1997) 'Genre analysis of article introductions in economics', *ASp. La Revue du Geras*, 15, 409-426.
- Lakic, I. (2010) 'Analysing genre: research article introductions in economics', *The Journal of Linguistic and Intercultural Education*, 3, 83-99.
- Lakić, I. (2015) 'The rhetorical structure of economics research article introductions and its syntactic and lexical exponents', in Lakić, I., Živković, B. and Vuković, M., eds., *Academic discourse across cultures*, Newcastle upon Tyne: Cambridge Scholars Publishing, 42-63.
- Laviosa, S. (2002) *Corpus-based translation studies: theory, findings, applications*, Amsterdam: Rodopi.

- Lee, D. (2001) 'Genres, registers, text types, domains, and styles: clarifying the concepts and navigating a path through the BNC jungle', *Language Learning and Technology*, 5(3), 37-72.
- Lee, D. and Swales, J. (2006) 'A corpus-based EAP course for NNS doctoral students: moving from available specialized corpora to self-compiled corpora', *English for Specific Purposes*, 25(1), 56-75.
- Lee, J.J. and Casal, J.E. (2014) 'Metadiscourse in results and discussion chapters: a cross-linguistic analysis of English and Spanish thesis writers in engineering', *System*, 46, 39-54.
- Lee, J.J. and Deakin, L. (2016) 'Interactions in L1 and L2 undergraduate student writing: interactional metadiscourse in successful and less-successful argumentative essays', *Journal of Second Language Writing*, 33, 21-34, available: <http://dx.doi.org/10.1016/j.jslw.2016.06.004>.
- Lee, N.I. (2011) 'A suggestion from research on contrastive rhetoric', *Proceedings of the International Conference: Doing Research in Applied Linguistics*, 33-42.
- Leech, G.N. (1983) *Principles of pragmatics*, New York: Longman.
- Lettrilliart, L. and Schott, A.M. (2007) 'Rédiger et publier un article de recherche en médecine générale', *La Revue du Praticien*, 21, 629-632.
- Levinson, S.C. (1983) *Pragmatics*, Cambridge: Cambridge University Press.
- Li, L.J. and Ge, G.C. (2009) 'Genre analysis: Structural and linguistic evolution of the English-medium medical research article (1985–2004)', *English for Specific Purposes*, 28(2), 93-104.
- Li, T. and Wharton, S. (2012) 'Metadiscourse repertoire of L1 Mandarin undergraduates writing in English: a cross-contextual, cross-disciplinary study', *Journal of English for Academic Purposes*, 11(4), 345-356, available: <http://dx.doi.org/10.1016/j.jeap.2012.07.004>.
- Li, Y. and Qian, D.D. (2010) 'Profiling the academic word list (AWL) in a financial corpus', *System*, 38(3), 402-411, available: <http://dx.doi.org/10.1016/j.system.2010.06.015>.
- Liardét, C.L. (2016) 'Grammatical metaphor: distinguishing success', *Journal of English for Academic Purposes*, 22, 109-118, available: <http://dx.doi.org/10.1016/j.jeap.2016.01.009>.

- Liddicoat, A. (1992) 'The use of the active and passive in French scientific prose: Some examples from the biological sciences', *Rassegna Italiana di Linguistica Applicata*, 14(2), 105-121.
- Liddicoat, A. (1993) 'Through the mirror darkly: epistemic modality in French scientific writing', Paper presented at the 18th ALAA National Congress, Adelaide, 20 September - 1 October.
- Liddicoat, A. (1995) 'Hedging in French scientific and technical writing', Paper presented at the 20th ALAA National Congress, Canberra, 19-22 September.
- Liddicoat, A. (1997) 'Communicating within cultures, communicating across cultures, communicating between cultures', in Golebiowski, Z. and Borland, H., eds., *Academic communication across disciplines and cultures*. Melbourne: Victoria University of Technology, 13-23.
- Liddicoat, A. (2005) 'Writing about knowing in science: aspects of hedging French scientific writing', *LSP and professional communication*, 5(2), 8-27.
- Lin, L. and Evans, S. (2012) 'Structural patterns in empirical research articles: a cross-disciplinary study', *English for Specific Purposes*, 31(3), 150-160.
- Lindquist, H. and Levin, M. (2000) 'Apples and oranges: on comparing data from different corpora', *Language and Computers*, 33, 201-214.
- Liu, L. and Li, L. (2016) 'Noun phrase complexity in EFL academic writing: a corpus-based study of postgraduate academic writing', *Journal of Asia TEFL*, 13(1), 48-65, available: <http://dx.doi.org/10.18823/asiatefl.2016.13.1.4.48>.
- Llorach Alarcos, E. (1999) *Gramática de la lengua española*, Madrid: Espasa Calpe.
- LoCastro, V. (2008) 'Long sentences and floating commas. Mexican students' rhetorical practices and the sociocultural context', in Connor, U., Nagelhout, E. and Rozycki, W., eds., *Contrastive rhetoric: reaching to intercultural rhetoric*, Amsterdam: John Benjamins, 195-217.
- Loffler-Laurian, A.M. (1980) 'L'Expression du locuteur dans les discours scientifiques', *Revue de Linguistique Romane*, 44, 135-157.
- Lorés-Sanz, R. (2006) '"I will argue that": first person pronouns as metadiscoursal devices in research article abstracts in English and Spanish', *ESP across Cultures*, 3, 23-40.

- Lorés-Sanz, R., (2011a) 'The construction of the author's voice in academic writing: the interplay of cultural and disciplinary factors', *Text & Talk*, 31(2), 173-193.
- Lorés-Sanz, R., (2011b) 'The study of authorial voice: using a Spanish-English corpus to explore linguistic transference', *Corpora*, 6(1), 1-24.
- Lorés-Sanz, R., Mur-Dueñas, P. and Lafuente-Millán, E. (2010) *Constructing interpersonal: multiple perspectives on written academic genres*, Newcastle upon Tyne: Cambridge Scholars Publisher.
- Love, R., Dembry, C., Hardie, A., Brezina, V., and McEnery, T. (2017) 'The Spoken BNC2014: designing and building a spoken corpus of everyday conversations', *International Journal of Corpus Linguistics*, 22(3), 319-344.
- Lu, P.Y. and Corbett, J. (2012) *English in medical education an intercultural approach to teaching language and values*, Bristol: Multilingual Matters.
- MacKenzie, I. (2015) 'Rethinking reader and writer responsibility in academic English', *Applied Linguistics Review*, 6(1), 1-21.
- Malcolm, L. (1987) 'What rules govern tense usage in scientific articles?', *English for Specific Purposes*, 6(1), 31-43.
- Marckwardt, A. (1972) 'Changing winds and shifting sands', *MST English Quarterly*, 21(3), 3-11.
- Martin, E. (2010) 'Designing and implementing a french-for-specific-purposes (FSP) program: lessons learned from ESP', *Global Business Languages*, 5(1), 3-21.
- Martin, J.R. (2001) 'Language genre and register', in Burns, A. and Coffin, C., eds., *Analysing English in a global context: a reader*, New York: Routledge, 149-166.
- Martin, J.R. (2009) 'Genre and language learning: a social semiotic perspective', *Linguistics and Education*, 20(1), 10-21, available: <http://dx.doi.org/10.1016/j.linged.2009.01.003>.
- Martin, J.R. and Rose, D. (2007) *Working with discourse: meaning beyond the clause*, London: Continuum.
- Martin, J.R. and White, P.R.R. (2005) *The language of evaluation: appraisal in English*, Basingstoke: Palgrave Macmillan.
- Martín, P. (2003) 'A genre analysis of English and Spanish research paper abstracts in experimental social sciences', *English for Specific Purposes*, 22(1), 25-43.

- Martín, P., Rey-Rocha, J., Burgess, S. and Moreno, A.I. (2014) 'Publishing research in English-language journals: attitudes, strategies and difficulties of multilingual scholars of medicine', *Journal of English for Academic Purposes*, 16, 57-67.
- Martínez, I.A. (2001) 'Impersonality in the research article as revealed by analysis of the transitivity structure', *English for Specific Purposes*, 20(3), 227-247.
- Marton, W. (1980) 'Equivalence and congruence in transformational contrastive studies', in Fisiak, J., ed., *Theoretical issues in contrastive linguistics*, Amsterdam: John Benjamins, 19-28.
- Marzá, N.E. (2014) 'A practical corpus-based approach to teaching English for tourism', *International Journal of Applied Linguistics and English Literature*, 3(1), 129-136, available: <http://dx.doi.org/10.7575/aiac.ijalel.v.3n.1p.129>.
- Marzo, S., Heylen, K. and De Sutter, G. (2012) 'Developments in corpus-based contrastive linguistics', in Marzo, S., Heylen, K. and De Sutter, G., eds., *Corpus studies in contrastive linguistics*, New York: Routledge, 3-13.
- Mauranen, A. (1993) *Cultural differences in academic rhetoric: a textlinguistic study*, New York: Peter Lang.
- Mauranen, A. and Bondi, M. (2003) 'Evaluative language use in academic discourse', *Journal of English for Academic Purposes*, 2(4), 269-271, available: [http://dx.doi.org/10.1016/S1475-1585\(03\)00045-6](http://dx.doi.org/10.1016/S1475-1585(03)00045-6).
- McCabe, A. and Heilman, K. (2007) 'Textual and interpersonal differences between a news report and an editorial', *Revista Alicantina de Estudios Ingleses*, 20, 139-156.
- McCarthy, M. (2016) 'Putting the CEFR to good use: designing grammars based on learner-corpus evidence', *Language Teaching*, 49(1), 99-115, available: <http://dx.doi.org/10.1017/S0261444813000189>.
- McCarthy, M. and O'Keeffe, A. (2010) 'Historical perspective: what are corpora and how have they evolved?', in O'Keeffe, A. and McCarthy, M., eds., *The Routledge handbook of corpus linguistics*, London: Routledge, 3-13.
- McCarthy, M., McCarten, J. and Sandiford, H. (2004-2006) *Touchstone (four levels)*, Cambridge: Cambridge University Press.
- McEnery, T. and Hardie, A. (2012) *Corpus linguistics: method, theory and practice*, Cambridge: Cambridge University Press.

- McEnery, T. and Wilson, A. (2001) *Corpus linguistics: an introduction*, Edinburgh: Edinburgh University Press.
- McEnery, A. and Xiao, R. (2005) 'Help or Help to: what do corpora have to say?', *English Studies*, 86(2), 161-187, available: <http://dx.doi.org/10.1080/0013838042000339880>.
- McEnery, T. and Xiao, R. (2008) 'Parallel and comparable corpora: what is happening?', in Anderman, G. A. and Rogers, M., eds., *Incorporating corpora: the linguist and the translator*, Clevedon: Multilingual Matters, 18-31.
- McEnery, T. and Xiao, R. (2010) *Corpus-based contrastive studies of English and Chinese*, London: Routledge.
- McEnery, T. and Xiao, R. (2011) 'What corpora can offer in language teaching and learning', in Hinkel, E., ed., *Handbook of research in second language teaching and learning*, London: Routledge, 364-380.
- McGrath, L. and Kuteeva, M. (2012) 'Stance and engagement in pure mathematics research articles: linking discourse features to disciplinary practices', *English for Specific Purposes*, 31(3), 161-173.
- McLeod, S. and Miraglia, E. (2001) 'Writing across the curriculum in a time of change', in McLeod, S., H., Miraglia, E., Soven, M. and Thaiss, C., eds., *WAC for the new millennium: strategies for continuing writing-across-the-curriculum programs*, Urbana: NCTE, 1-27.
- Miller, C.R. (2003) 'Genre as social action', in Freedman, A. and Medway, P., eds., *Genre in the new rhetoric*, London: Taylor and Francis, 20-36.
- Ministerio de Educación (2016) *Teseo*, available: <https://www.educacion.gob.es/teseo/irGestionarConsulta.do> [accessed 09 Nov 2016].
- Mishan, F. (2005) *Designing authenticity into language learning materials*, Bristol: Intellect.
- Mobjörk, M. (2010) 'Consulting versus participatory transdisciplinarity: a refined classification of transdisciplinary research', *Futures*, 42(8), 866-873.
- Montemayor-Borsinger, A. (2007) 'El análisis de la organización del discurso literario en español. Una propuesta desde la lingüística sistémico funcional', *Co-herencia*, 4(7), 133-153.

- Montgomery, S. (1996) *The scientific voice*, New York: The Guildford Press.
- Moreno, A.I. (2004) 'Retrospective labelling in premise–conclusion metatext: an English–Spanish contrastive study of research articles on business and economics', *Journal of English for Academic Purposes*, 3(4), 321-339, available: <http://dx.doi.org/10.1016/j.jeap.2004.07.005>.
- Moreno, A.I. (2008) 'The importance of comparable corpora in cross-cultural studies', in Connor, U., Nagelhout, E. and Rozycki, W., eds., *Contrastive rhetoric: reaching to intercultural rhetoric* Amsterdam: John Benjamins, 25-41.
- Morton, J. (2016) '“Adjacent worlds”: an analysis of a genre at the intersection of academic and professional communities', *Journal of English for Academic Purposes*, 22, 54-63, available: <http://dx.doi.org/10.1016/j.jeap.2016.01.003>.
- Mufwene, S.S. (2005) 'Globalization and the myth of killer languages: what's really going on', *Perspectives on Endangerment*, 5, 19-48.
- Mullally, G., Byrne, E. and Sage, C. (2016) 'Disciplines, perspectives and conversations', in Byrne, E., Mullally, G. and Sage, C., eds., *Transdisciplinary perspectives on transitions to sustainability*, 21-40.
- Mur-Dueñas, P.M. (2007a) 'Same genre, same discipline; however, there are differences: a cross-cultural analysis of logical markers in academic writing', *ESP across Cultures*, 4, 37-53.
- Mur-Dueñas, P.M. (2007b) 'I/we focus on...: a cross-cultural analysis of self-mentions in business management research articles', *Journal of English for Academic Purposes*, 6, 143-162.
- Mur-Dueñas, P. (2011) 'An intercultural analysis of metadiscourse features in research articles written in English and in Spanish', *Journal of Pragmatics*, 43(12), 3068-3079, available: <http://dx.doi.org/10.1016/j.pragma.2011.05.002>.
- Nagano, R. (2010) 'Question forms in journal article titles', *Proceedings of the 4th Doctoral Conference in Applied Linguistics*, Budapest: Research Institute for Linguistics of the Hungarian Academy of Sciences.
- Nagano, R. (2015) 'Research article titles and disciplinary conventions: a corpus study of eight disciplines', *Journal of Academic Writing*, 5(1), 133-144.
- Nathan, P. (2013) 'Academic writing in the business school: the genre of the business case report', *Journal of English for Academic Purposes*, 12(1), 57-68.

- Nesi, H., Gardner, S., and Thompson, P. (2008), *British Academic Written English Corpus*, Oxford Text Archive.
- Nesselhauf, N. (2006) 'Learner corpora and their potential for language teaching', in Sinclair, J.M., ed., *How to use corpora in language teaching*, Amsterdam: John Benjamins.
- Neubert, A. (1994) 'Competence in translation: a complex skill, how to study and how to teach it', in Snell-Hornby, M., Pöschhacker, F. and Kaindl, K., eds., *Translation studies: an interdisciplinary*, Amsterdam: John Benjamins, 411-420.
- Nichols, J. (1988) 'Nominalization and assertion in scientific Russian prose', in Haiman, J. and Thompson, S., eds., *Clause combining in grammar and discourse*, Amsterdam: John Benjamins, 399-428.
- Nølke, H., Fløttum, K. and Norén, C. (2004) *Scapoline: la théorie scandinave de la polyphonie linguistique*, Paris: Kimé.
- Northedge, A. (2005) *The good study guide*, Oxford: Open University Press.
- O'Keeffe, A. and Farr, F. (2012) 'Using language corpora in initial teacher education: pedagogic issues and practical applications', in Biber, D. and Reppen, R., eds., *Corpus linguistics vol. 4*, Los Angeles: Sage, 335-362.
- O'Keeffe, A. and Mark, G. (2017) 'The English grammar profile of learner competence', *International Journal of Corpus Linguistics*, 22(4), 457-489.
- O'Keeffe, A., McCarthy, M. and Carter, R. (2007) *From corpus to classroom: language use and language teaching*, Cambridge: Cambridge University Press.
- O'Sullivan, Í. (2010) 'Using corpora to enhance learners' academic writing skills in French', *Revue Francaise de Linguistique Appliquée*, 15(2), 21-35.
- Oleksy, W. (1986) 'Some recent approaches to equivalence in contrastive studies', in Fisiak, J., Kastovsky, D. and Szwedek, A. J., eds., *Linguistics across historical and geographical boundaries: in honour of Jacek Fisiak on the occasion of his fiftieth birthday*, Berlin: De Gruyter, 1405-1420.
- Oliver del Olmo, S. (2004) *Análisis contrastivo español/inglés de la atenuación retórica en el discurso médico*, unpublished thesis, Universitat Pompeu Fabra.
- Oliver del Olmo, S. (2005) 'The role of passive voice in hedging medical discourse: a corpus-based study on English and Spanish research articles', *Revista De Lengua Para Fines Específicos*, 12, 205-218.

- Oliver del Olmo, S. (2015) 'Spanish authors dealing with hedging or the challenges of scholarly publication in English L2', in Plo, A.R. and Pérez-Llantada, C., eds., *Volume 2 English as a scientific and research language. Debates and discourses*. Berlin: De Gruyter, 141-157.
- Oller, J.W. and Ziahosseiny, S.M. (1970) 'The contrastive analysis hypothesis and spelling errors', *Language Learning*, 20(2), 183-189, available: <http://dx.doi.org/10.1111/j.1467-1770.1970.tb00475.x>.
- Orta, I.V. (2010) 'A contrastive analysis of the use of modal verbs in the expression of epistemic stance in business management research articles in English and Spanish', *Ibérica, Revista de la Asociación Europea de Lenguas para Fines Específicos*, 19, 77-95.
- Orta, I.V., Millán, E.L., Sanz, R.L. and Dueñas, M.P.M. (2006) 'How to explore academic writing from metadiscourse as an integrated framework of interpersonal meaning: three perspectives of analysis', *Proceedings of the 5th International AELFE Conference*, Prensas Universitarias de Zaragoza, 197-208.
- Oster, S. (1981) 'The use of tenses in "reporting past literature" in EST', in Selinker, L., Tarone, E. and Hanzeli, A., eds., *English for academic and technical purposes*, Rowley: Newbury House, 76-90.
- Paré, A. (2014) 'Rhetorical genre theory and academic literacy', *Journal of Academic Language and Learning*, 8(1), 83-94.
- Parkinson, J. and Adendorff, R. (2005) 'Variable discursive constructions of three genres of science', *Southern African Linguistics and Applied Language Studies*, 23(3), 281-303.
- Pennycook, A. (1997) 'Vulgar pragmatism, critical pragmatism, and EAP', *English for Specific Purposes*, 16, 253-269.
- Pennycook, A. (2001) *Critical applied linguistics: a critical introduction*. Mahwah: Erlbaum.
- Pérez-Llantada, C. (2010) 'The discourse functions of metadiscourse in published academic writing: Issues of culture and language', *Nordic Journal of English Studies*, 9(2), 41-68.
- Pérez-Llantada, Plo, A.R. and Ferguson, G. (2011) "'You don't say what you know, only what you can": the perceptions and practices of senior Spanish academics

- regarding research dissemination in English', *English for Specific Purposes*, 30(1), 18-30.
- Pérez, M.D.M.S. and Ramiro, M.S.S. (2015) 'Writing professional genres in a second language: results from a Spanish university CLIL context', *American Journal of Educational Research*, 3(5), 576-580.
- Pérez-Paredes, P. and Díez-Bedmar, B. (2012) 'The use of intensifying adverbs in learner writing', in Tono, Y., Kawaguchi, Y. and Minegishi, M., eds., *Developmental and crosslinguistic perspectives in learner corpus research*. Amsterdam: John Benjamins, 105-124.
- Phillipson, R. (2006) 'English, a cuckoo in the European higher education nest of languages?', *European Journal of English Studies*, 10(1), 13-32.
- Phillipson, R. (2008a) 'Lingua franca or lingua frankensteinia? English in European integration and globalisation', *World Englishes*, 27, 250-267.
- Phillipson, R. (2008b) 'The linguistic imperialism of neoliberal empire', *Critical Inquiry in Language Studies*, 5(1), 1-43.
- Pic, E. and Furmaniak, G. (2014) 'Questioning certainty in research articles and popular science articles. A case-study of modalized wh-interrogatives', in Zuczkowski, A. and Bongelli, R., eds., *Communicating certainty and uncertainty in medical, supportive and scientific contexts*, Amsterdam: John Benjamins, 371-389.
- Piotrowski, T. (1994) *Problems in bilingual lexicography*, Wrocław: Wydawnictwo Uniwersytetu Wrocławskiego.
- Pisanski-Peterlin, A.P. (2005) 'Text organizing metatext in research articles: an English-Slovene contrastive analysis', *English for Specific Purposes*, 24(3), 307-319.
- Plann, S. (1982) 'Indirect Questions in Spanish', *Linguistic Inquiry*, 13(2), 297-312.
- Pöldvere, N., Fuoli, M. and Paradis, C. (2016) 'A study of dialogic expansion and contraction in spoken discourse using corpus and experimental techniques', *Corpora*, 11(2), 191-225, available: <http://dx.doi.org/10.3366/cor.2016.0092>.
- Poole, R. (2016) 'A corpus-aided approach for the teaching and learning of rhetoric in an undergraduate composition course for L2 writers', *Journal of English for Academic Purposes*, 21, 99-109, available: <http://dx.doi.org/10.1016/j.jeap.2015.12.003>.

- Pountain, C.J. and Kattán-Ibarra, J. (2004) *Modern Spanish grammar: a practical guide*, London: Routledge.
- Price, L. and Robinson, L. (2016) ‘‘Being in a knowledge space’’: information behaviour of cult media fan communities’, *Journal of Information Science*, 43(5), 649-664, available: <http://dx.doi.org/10.1177/0165551516658821>.
- Prior, P. (1998) *Writing/disciplinarity: a sociohistoric account of literate activity in the academy*, Mahwah: Erlbaum.
- Prior, P. (2003) ‘Are communities of practice really an alternative to discourse communities?’ Paper presented at the in the American Association of Applied Linguistics, Arlington, Virginia, 22-25 March.
- Rabadán, R. (2005) ‘Tools for English-Spanish cross-linguistic applied research’, *Journal of English Studies*, (5), 309-324.
- Rabadán, R. (2011) ‘Any into Spanish: a corpus-based analysis of a translation problem’, *Linguistica Pragmática*, 21(2), 57-69, available: <http://dx.doi.org/10.2478/v10017-011-0005-y>.
- Rabadán, R. and Izquierdo, M. (2013) *A corpus-based study of English approximate negators and their translations into Spanish*, *Bergen Language and Linguistics Studies*, 3(1), 43-60.
- Rabatel, A. (1998) *La construction textuelle du point de vue*, Lausanne: Delachaux et Niestlé.
- Rahimivand, M. and Kuhi, D. (2014) ‘An exploration of discoursal construction of identity in academic writin’, *Procedia - Social and Behavioral Sciences*, 98, 1492-1501, available: <http://dx.doi.org/10.1016/j.sbspro.2014.03.570>.
- Ramos, I.D.D. (2015) ‘Error, register, grammatical-rhetorical, and genre analyses on nursing students’ case studies and thesis proposals’, *The Asian ESP Journal*, 2(2), 62-97.
- Read, J. and Carroll, J. (2012) ‘Annotating expressions of appraisal in English’, *Language Resources and Evaluation*, 46(3), 421-447, available: <http://dx.doi.org/10.1007/s10579-010-9135-7>.
- Rentel, N. (2006) ‘Evaluation in Italian and French research articles in linguistics’, in Suomela-Salmi, E. and Dervin, F., eds., *Perspectives inter-culturelles et inter-*

- linguistiques sur le discours académique Cross-cultural and Cross-linguistic Perspectives on Academic Discourse*, Turku: University of Turku, 59-70.
- Revilla, S. (1980) *Gramática española moderna*, México: McGraw-Hill.
- Rinck, F. (2011) 'Former à (et par) l'écrit de recherche. Quels enjeux, quelles exigences?', *Le français aujourd'hui*, (3), 79-89.
- Rodgers, O., Chambers, A. and LeBaron-Earle, F. (2011) 'Corpora in the LSP classroom: A learner-centred corpus of French for biotechnologists', *International Journal of Corpus Linguistics*, 16(3), 391-411, available: <http://dx.doi.org/10.1075/ijcl.16.3.06rod>.
- Rohdenburg, G. (1992) 'Bermerkungen zu infinite konstuktionen im Englischen und Deutschen', in Mair, C. and Markus, M., eds., *New departures in contrastive linguistics*, Innsbruck: University of Innsbruck Press, 187-207.
- Römer, U. (2011) 'Corpus research applications in second language teaching', *Annual Review of Applied Linguistics*, 31, 205-225.
- Romero, M. (2005) 'Concealed questions and specificational subjects', *Linguistics and Philosophy*, 28(6), 687-737.
- Rongen Breivega, K., Dahl, T. and Fløttum, K. (2002) 'Traces of self and others in research articles. A comparative pilot study of English, French and Norwegian research articles in medicine, economics and linguistics', *International Journal of Applied Linguistics*, 12(2), 218-239, available: <http://dx.doi.org/10.1111/1473-4192.00032>.
- Rowley-Jolivet, E. and Carter-Thomas, S. (2005a) 'Genre awareness and rhetorical appropriacy: manipulation of information structure by NS and NNS scientists in the international conference settings', *English for Specific Purposes*, 24 (1), 41-64.
- Rowley-Jolivet, E. and Carter-Thomas, S. (2005b) 'The rhetoric of conference presentation introductions: context, argument and interaction', *International Journal of Applied Linguistics*, 15 (1), 45-71.
- Rueda, A.M. and Alamán, A.P. (2015) 'Syntactic and discursive values of Twitter hashtags', *Circulo de Linguistica Aplicada a la Comunicacion*, 64, 58-83, available: <http://dx.doi.org/10.5209/rev-CLAC.2015.v64.51278>.

- Ruegg, R. and Sugiyama, Y. (2013) 'Organization of ideas in writing: what are raters sensitive to?', *Language Testing in Asia*, 3(1), 1-13.
- Russell, D.R., Lea, M., Parker, J., Street, B. and Donahue, T. (2009) 'Exploring notions of genre in 'academic literacies' and 'writing across the curriculum': approaches across countries and contexts', in Bazerman, C., Bonini, A. and Figueiredo, D., eds., *Genre in a changing world. Perspectives on writing*, Colorado: Parlor Press, 459-491.
- Ryvitytė, B. (2003) 'Research article introductions: variations across disciplines and cultures', *Kalbotyra*, 53, 93-100.
- Sacks, H. (1984) 'Methodological remarks', in Maxwell, J. and Heritage, J., eds., *Structures of social action: studies in conversation analysis*, Cambridge: Cambridge University Press, 21-27.
- Sampson, G. (2013) 'The empirical trend: ten years on', *International Journal of Corpus Linguistics*, 18(2), 281-289, available: <http://dx.doi.org/10.1075/ijcl.18.2.05sam>.
- Salager-Meyer, F. (2000) 'Rhetorical evolution of oppositional discourse in French academic writing: Oppositional discourse in academic writing', *Hermes*, 25, 23-48.
- Salager-Meyer, F. (2011a) 'Scientific discourse and contrastive linguistics: hedging' *European Science Editing*, 37(2), 35-37.
- Salager-Meyer, F. (2011b) 'Scientific discourse and contrastive linguistics: explicitness and the concept of reader/writer responsible languages', *European Science Editing*, 37(3), 71-72.
- Salager-Meyer, F. and Ariza, M.A. (2004) 'Negative appraisals in academic book reviews: a cross-linguistic approach', *Intercultural aspects of specialized communication*, 149-172.
- Salager-Meyer, F., Ariza, M.A. and Zambrano, N. (2003) 'The scimitar, the dagger and the glove: intercultural differences in the rhetoric of criticism in Spanish, French and English medical discourse (1930-1995)', *English for Specific Purposes*, 22(3), 223-247.
- Salas, M.D. (2015) 'Reflexive metadiscourse in research articles in Spanish: variation across three disciplines (linguistics, economics and medicine)', *Journal of*

- Salazar, D., Ventura, A. and Verdaguer, I. (2013) 'A cross-disciplinary analysis of personal and impersonal features in English and Spanish scientific writing', in Verdaguer, I, Laso, N.J. and Salazar, D., eds., *Biomedical English: a corpus-based approach*, Amsterdam: John Benjamins, 121-144.
- San Roque, L., 2016. "'Where' questions and their responses in Duna (Papua New Guinea)", *Open Linguistics*, 2(1), 85-104.
- Sánchez-López, L. (2012) 'Spanish for specific purposes', in Chapelle, C.A., ed., *The Encyclopedia of Applied Linguistics*, Oxford: Wiley-Blackwell, 1-18.
- Santos, J. and da Silva, P. (2016) 'Issues with publishing abstracts in English: challenges for Portuguese linguists' authorial voices', *Publications*, 4(2), 1-17.
- Sawyer, A.G., Laran, J. and Xu, J. (2008) 'The readability of marketing journals: are award-winning articles better written?', *Journal of Marketing*, 72(1), 108-117.
- Šćepanović, A. (2015) 'Genre characteristics of research article introductions in civil engineering', in Lakić, I., Živković, B. and Vuković, M., eds., *Academic discourse across cultures*, Newcastle upon Tyne, UK: Cambridge Scholars Publishing, 64-84.
- Schmied, J. (2011) 'Teaching and learning contrastive linguistics using an EU translation corpus with English, German, French and Spanish', in Kübler, N., ed., *Corpora, language, teaching, and resources: from theory to practice*, Bern: Peter Lang, 165-184.
- Schluer, J. (2015) 'English as a lingua franca in linguistics? A case study of German linguists' language use in publications', in Plo, A.R. and Pérez-Llantada, C., eds., *Volume 2 English as a scientific and research language. Debates and discourses*. Berlin: De Gruyter, 233-260.
- Schulze, J. (2009) 'Testing: a systemic functional view of high stakes test preparation materials', *Colombian Applied Linguistics Journal*, 11, 80-92.
- Sergeant, P. (2012) 'Disciplinarity and the study of world Englishes', *World Englishes*, 31(1), 113-129.
- Searle, J.R. (1969) *Speech acts: an essay in the philosophy of language*, Cambridge: Cambridge University Press.

- Searle, J.R., Willis, S. and Vanderveken, D. (1985) *Foundations of illocutionary logic*, Cambridge: Cambridge University Press.
- Seidlhofer, B. (2010) 'Lingua franca English: the European context', in Kirkpatrick, A. ed., *The Routledge handbook of world Englishes*, London: Routledge, 355–371.
- Selinker, L. (1972) 'Interlanguage', *International Review of Applied Linguistics in Language Teaching*, 10(4), 209-232.
- Senft, G. (2014) *Understanding pragmatics*, Abingdon: Routledge.
- Shah, D.V., Cappella, J.N., and Neuman, W.R. (2015) 'Big data, digital media, and computational social science: possibilities and perils', *The Annals of the American Academy of Political and Social Science*, 659(1), 6-13, <https://doi.org/10.1177/0002716215572084>.
- Shaw, P. (1992) 'Reasons for the correlation of voice, tense, and sentence function in reporting verbs', *Applied Linguistics*, 13(3), 302-319.
- Shaw, P. (2003) 'Evaluation and promotion across languages', *Journal of English for Academic Purposes*, 2(4), 343-357, available: [http://dx.doi.org/10.1016/S1475-1585\(03\)00050-X](http://dx.doi.org/10.1016/S1475-1585(03)00050-X).
- Shi, L. (2002) 'How western-trained Chinese tesol professionals publish in their home environment', *TESOL Quarterly*, 36(4), 625-634, available: <http://dx.doi.org/10.2307/3588245>.
- Siepmann, D. (2006) 'Academic writing and culture: an overview of differences between English, French and German', *Meta: Journal des traducteurs Meta*, 51(1), 131-150.
- Siewierska, A. (1986) *The passive: a comparative linguistic analysis*, London: Croom Helm.
- Simpson, J. (2011) 'Introduction', in Simpson, J., ed., *The Routledge handbook of applied linguistics*, London: Routledge, 1-9.
- Sketch Engine. (2019) Sketch Engine [Computer Software], Brno: Brno University of Technology. Available from <https://www.sketchengine.eu/>.
- Smith, D.A. (1987) 'Variation in field related genres', *ELR Journal*, 1. 10-32.
- Smith, N.V. and Allott, N. (2016) *Chomsky: ideas and ideals*, Cambridge: Cambridge University Press.

- Smith, V. (2007) 'Speaking the same language. dialogue in academic writing', in Delanoy, W., Helbig, J.R. and James, A., eds., *Towards a dialogic anglistics*, London: Lit, 160-176.
- Soler, V. (2007) 'Writing titles in science: an exploratory study', *English for Specific Purposes*, 26, 90-102.
- Soler, V. (2009) 'Títulos científicos en lengua española: estudio exploratorio', *Lebende Sprachen*, 54(2), 50-58.
- Soler, V. (2011) 'Comparative and contrastive observations on scientific titles written in English and Spanish', *English for Specific Purposes*, 30(2), 124-137.
- Suomela-Salmi, E. and Dervin, F. eds. (2011) *Perspectives interculturelles et interlinguistiques sur le discours académique. Volume 1*, Turku: University of Turku.
- Spack, R. (1997) 'The rhetorical construction of multilingual students', *Tesol Quarterly*, 31(4), 765-774.
- Sridhar, S. (1975) 'Contrastive analysis, error analysis and interlanguage: three phases of one goal?', *Studies in language learning*, 1, 1-35.
- Strowe, A. (2011) 'Is simpatico possible in translation?: The 1620 translation of the Decameron and the case for similarity', *Translator*, 17, 51-75.
- Suarez Tejerina, L. (2005) 'Is evaluation structure bound? An English-Spanish contrastive study of book reviews', in Tognini-Bonelli, E. and Del Lungo Camiciotti, G., eds., *Strategies in academic discourse*, Amsterdam: John Benjamins, 117-132.
- Suñer, M. (1993) 'About indirect questions and semi-questions', *Linguistics and Philosophy*, 16(1), 45-77.
- Sutton, C. (1995) 'Quelques questions sur l'écriture et la science: une vue personnelle d'outre-Manche. Repères', *Recherches en Didactique Dd Français Langue Maternelle*, 12(1), 37-52.
- Swales, J.M. (1988) 'Discourse communities, genres and English as an international language', *World Englishes*, 7(2), 211-220.
- Swales, J.M. (1990) *Genre analysis: English in academic and research settings*, Cambridge: Cambridge University Press.

- Swales, J.M. (1993) 'Genre and Engagement', *Revue Belge de Philologie et d'Histoire*, 71(3), 687-698.
- Swales, J.M. (1997) 'English as tyrannosaurus rex', *World Englishes*, 16, 373-382.
- Swales, J.M. (1998) *Other floors, other voices: a textography of a small university building*, Mahwah: Erlbaum.
- Swales, J.M. (2004) *Research genres: explorations and applications*, Cambridge: Cambridge University Press.
- Swales, J.M. (2016) 'Reflections on the concept of discourse community', *ASp. La Revue du GERAS*, (69), 7-19.
- Swales, J.M. and Feak, C.B. (1994) *Academic writing for graduate students: essential tasks and skills*, Ann Arbor: University of Michigan Press.
- Swales, J.M. and Feak, C.B. (2000) *English in today's research world*, Ann Arbor: University of Michigan Press.
- Swales, J.M. and Feak, C.B. (2012) *Academic writing for graduate students: Essential tasks and skills* (3e), Ann Arbor: University of Michigan Press.
- Syal, P. and Jindal, D. (2007) *An introduction to linguistics*, New Delhi: Prentice-Hall of India.
- Taft, M., Kacanas, D., Huen, W. and Chan, R. (2011) 'An empirical demonstration of contrastive rhetoric: preference for rhetorical structure depends on one's first language', *Intercultural Pragmatics*, 8(4), 503-516.
- Tangpijaikul, M. (2014) 'Preparing business vocabulary for the ESP classroom', *RELC Journal*, 45(1), 51-65, available: <http://dx.doi.org/10.1177/0033688214522641>.
- Tarone, E., Dwyer, S., Gillette, S. and Icke, V. (1981) 'On the use of the passive in two astrophysics journal papers', *English for Specific Purposes*, 1(2), 123-140.
- Taylor, C. and Marchi, A. (2018) 'Introduction: partiality and reflexivity', in Taylor, C. and Marchi, A. eds., *Corpus approaches to discourse: a critical review* London: Routledge, 1-17.
- Teich, E. (2003) *Crosslinguistic variation in system and text: a methodology for the investigation of translations and comparable texts*, Berlin: De Gruyter.
- Thomas, J. (2015) *Discovering English with Sketch Engine: a corpus-based approach*, Brno: Versatile.

- Thompson, G. (2001a) 'Interaction in academic writing: learning to argue with the reader', *Applied Linguistics*, 22(1), 58-78.
- Thompson, P. (2001b) *A pedagogically-motivated corpus-based examination of PhD theses: macrostructure, citation practices and uses of modal verbs*, unpublished thesis, University of Reading.
- Thompson, P. (2009) 'Shared disciplinary norms and individual traits in the writing of British undergraduates', in Gotti, M., ed., *Commonality and individuality in academic discourse*, New York: Peter Lang, 53-82.
- Tijdkink, J.K., Verbeke, R. and Smulders, Y.M. (2014) 'Publication pressure and scientific misconduct in medical scientists', *Journal of Empirical Research on Human Research Ethics*, 9(5), 64-71, available: <http://dx.doi.org/10.1177/1556264614552421>.
- Tijdkink, J.K., Vergouwen, A.C.M. and Smulders, Y.M. (2013) 'Publication pressure and burn out among Dutch medical professors: a nationwide survey', *PLoS ONE*, 8(9), available: <http://dx.doi.org/10.1371/journal.pone.0073381>.
- Timmis, I. (2015) *Corpus linguistics for ELT: research and practice*, London: Routledge.
- Tognini-Bonelli, E. (2001) *Corpus linguistics at work*, Amsterdam: John Benjamins.
- Tse, P. and Hyland, K. (2006) "'So what is the problem this book addresses?': Interactions in academic book reviews", *Text & Talk*, 26(6), 767-790.
- Tsui, A.B.M. (2006) 'What teachers have always wanted to know – and how corpora can help', in Sinclair, J.M., ed., *How to use corpora in language teaching*, Amsterdam: John Benjamins, 39-61.
- Tuck, J. (2016) "'That ain't going to get you a professorship': discourses of writing and the positioning of academics' work with student writers in UK higher education", *Studies in Higher Education*, 41(9), 1612-1626, available: <http://dx.doi.org/10.1080/03075079.2014.999320>.
- Tutin, A. (2010a) "'Dans cet article, nous souhaitons montrer que..." Lexique verbal et positionnement de l'auteur dans les articles en sciences humaines', *Lidil*, 41, 15-40.
- Tutin, A. (2010b) 'Evaluative adjectives in academic writing in the humanities and social sciences', in Lorés-Sanz, R., Mur-Dueñas, P. and Lafuente-Millán, E., eds.,

- Constructing interpersonal: multiple perspectives on written academic genres*, Newcastle upon Tyne: Cambridge Scholars Publishing, 219-242.
- Tutin, A. (2013) 'La phraséologie transdisciplinaire des écrits scientifiques: des collocations aux routines sémantico-rhétoriques', in Tutin, A. and Grossmann, F., eds., *L'écrit scientifique: du lexique au discours. Autour de scientext*, Rennes: Presses Universitaires de Rennes, 27-43.
- Uzuner, S. (2008) 'Multilingual scholars' participation in core/global academic communities: a literature review', *Journal of English for Academic Purposes*, 7(4), 250-263.
- Valero-Garcés, C. (1996) 'Contrastive ESP rhetoric: metatext in Spanish-English economics texts', *English for Specific Purposes*, 15(4), 279-294, available: [http://dx.doi.org/10.1016/S0889-4906\(96\)00013-0](http://dx.doi.org/10.1016/S0889-4906(96)00013-0).
- Van Bonn, S. and Swales, J.M. (2007) 'English and French journal abstracts in the language sciences: three exploratory studies', *Journal of English for Academic Purposes*, 6(2), 93-108, available: <http://dx.doi.org/10.1016/j.jeap.2007.04.001>.
- Van der Auwera, J. (2012) 'From contrastive linguistics to linguistic typology', *Languages in Contrast*, 12(1), 69-86.
- Vande Kopple, W. (2002) 'Metadiscourse, discourse, and issues in composition and rhetoric', in Barton, F. and Stygall, C., eds., *Discourse studies in composition*, Cresskill: Hampton Press, 91-113.
- Vanderveken, D. (2001) 'Illocutionary logic and discourse typology', *Revue Internationale de Philosophie*, (2), 243-255.
- Vassileva, I. (2000) *Who is the author?: a contrastive analysis of authorial presence in English, German, French, Russian, and Bulgarian academic discourse*, Sankt Augustin: Asgard.
- Vázquez Orta, I. and Giner, D. (2008) 'Beyond mood and modality: epistemic modality markers as hedges in research articles. A cross-disciplinary study', *Revista Alicantina de Estudios Ingleses*, 21, 171-190.
- Vold, E.T. (2006) 'Epistemic modality markers in research articles: a cross-linguistic and cross-disciplinary study', *International Journal of Applied Linguistics*, 16(1), 61-87.
- Wallwork, A. (2011) *English for writing research papers*, Heidelberg: Springer.

- Walsh, S. (2013) 'Corpus linguistics and conversation analysis at the interface: theoretical perspectives, practical outcomes', in Romero-Trillo, J., ed., *Yearbook of corpus linguistics and pragmatics 2013: new domains and methodologies*, Heidelberg: Springer, 37-53.
- Wardhaugh, R. (1970) 'The contrastive analysis hypothesis', *TESOL Quarterly*, 4(2), 123-130, available: <http://dx.doi.org/10.2307/3586182>.
- Webber, P. (1994) 'The function of questions in different medical English genres', *English for Specific Purpose*, 13, 257-68.
- Webster, J.J. (2009) 'An introduction to continuum companion to systemic functional linguistics', in Halliday, M.A.K. and Webster, J.J., eds., *Continuum companion to systemic functional linguistics*, London: Continuum, 1-12.
- Wendland, E. (2007) 'What's the difference? Similarity (and dissimilarity) from cross cultural perspective: some reflections upon the notions of "acceptability" in bible translation', in Arduini, S. and Hodgson, R., eds., *Similarity and difference in translation: proceedings of the international conference on similarity and translation*, Roma: Edizioni di storia e letteratura, 329-358.
- Wiegand, H.E. (2005) 'Äquivalenz, äquivalentendifferenzierung und Äquivalentpräsentation in zweisprachigen Wörterbüchern. Eine neue einheitliche Konzeption', in Pedersen, V., in Zettersten, V.H., and Mogensen, J.E., eds., *International symposium on, lexicography*, Tübingen: Max Niemeyer, 17-57.
- Wierzbicka, A. (1999) *Emotions across languages and cultures: diversity and universals*, Cambridge: Cambridge University Press.
- Willis, J. (1998) 'Concordances in the classroom without a computer: assembling and exploiting concordances of common words', in Tomlinson, B., ed., *Materials development in language teaching*, Cambridge: Cambridge University Press, 44-66.
- Winkler, A.C. and Methereil, J.R. (2012) *Writing the research paper: a handbook (8e)*, Boston: Wadsworth.
- Wojciechowska, S. (2016) 'Collocations and colligations of the metonymic shipment: exemplification in learners' dictionaries vs. corpus evidence', *International Journal of Lexicography*, 28(4), 465-489, available: <http://dx.doi.org/10.1093/ijl/ecv007>.

- Xie, J. (2016) 'Direct or indirect? Critical or uncritical? Evaluation in Chinese English-major MA thesis literature reviews', *Journal of English for Academic Purposes*, 23, 1-15, available: <http://dx.doi.org/10.1016/j.jeap.2016.05.001>.
- Yngve, V. and Wasik, Z. (2006) *Hard-science linguistics*, New York: Continuum.
- Yunick, S. (1997) 'Genres, registers and sociolinguistics', *World Englishes*, 16(3), 321-336.
- Zgusta, L. and Hartmann, R.R.K. (1984) 'Translational equivalence and the bilingual dictionary', in *LEXeter'83 Proceedings*, Tübingen: Max Niemeyer, 147-154.
- Zhonghua, L.I. (2015) 'Discipline-based English teaching from the perspective of SFL register theory', *Foreign Language World*, 3, 1-5.

Appendix

Appendix A: Articles in specon subcorpus

specon01

García de la Cruz, J. M. and Sánchez Díez, Á. (2001) 'Globalidad y concentración económica: implicaciones para la política económica y del desarrollo', *Revista de Economía Mundial*, 55(2), 243-262.

specon02

Pajares, E.M. and Molina, M.D.M.H. (2000) 'Cambios en las pautas de comercio exterior de los países del Magreb (1985-1995)', *Revista de Economía Mundial*, 2, 121-136.

specon03

Fernández, F. G. and de Azagra, L. M. (2001) 'Evolución y desarrollo crítico del sistema capitalista', *Revista de Economía Mundial*, 5, 167-181.

specon04

De Pablos, M.I.H. and Blanco, B.O. (2002) 'La política estructural comunitaria y la iniciativa comunitaria INTERREG en el horizonte de la ampliación de la Unión Europea', *Revista de Economía Mundial*, 6, 95-109.

specon05

García Mezquita, Y. and Sánchez Díez, Á. (2003) 'La internacionalización del sector energético español: la experiencia hacia América latina', *Revista de Economía Mundial*, 9, 53-79.

specon06

Romero, C.A. and Petrecolli, D. (2003) 'Desempeño, crisis y reformas en el sector eléctrico argentino: lecciones para países en desarrollo', *Revista de Economía Mundial*, 9, 21-50.

specon07

Fernández-Muro, C.G. and Aldasoro, I.O. (2003) 'Alcance versus desempeño financiero en los programas de microfinanzas: apuntes teóricos y algunos ejemplos', *Revista de Economía Mundial*, 9, 129-152.

specon08

Hornero, A.C. (2000) 'Organismos financieros internacionales: bancos regionales de desarrollo e instituciones financieras multilaterales' *Revista de Economía Mundial*, 3, 79-97.

specon09

Calderón, M.C.L. (2001) 'Los microcréditos: un nuevo instrumento de financiación para luchar contra la pobreza', *Revista de Economía Mundial*, 5, 121-138.

specon10

Fernández, S.G. (2002) 'La necesaria reforma institucional en la Unión Europea', *Revista de Economía Mundial*, 6, 49-62.

specon11

Salto, M.G. (2002) 'La Promoción de la actividad turística: ¿una política de desarrollo acertada?' *Revista de Economía Mundial*, 7, 111-113.

specon12

Miranda, M.J.V. (2003) 'Impactos sociales del proceso de transición económica de Europa del este', *Revista de Economía Mundial*, 8, 175-204.

specon13

Mier, M.Á.D. (2000) 'Un siglo de relaciones comerciales internacionales', *Revista de Economía Mundial*, 3, 11-47.

specon14

Vendrell, A.S. (2001) 'El papel de las ONGDs en la cooperación al desarrollo', *Revista de Economía Mundial*, 5, 85-97.

specon15

Solís, J.A.N. (2002) 'El debate sobre el sistema de financiación de la UE', *Revista de Economía Mundial*, 6, 63-84.

specon16

Anguita, J.P. (2003) 'La crisis económica de Argentina y la viabilidad de los sistemas de convertibilidad', *Revista de Economía Mundial*, 9, 111-128.

specon17

De Quinto Romero, J. (2003) 'Condicionantes de la estructura industrial en el sector energético de los países de la UE', *Revista de Economía Mundial*, 9, 41-52.

specon18

Drucker, A.G., Gómez González, V. and Magaña Rueda, S. (1999). '¿Contaminante o bien? La importancia del valor de los subproductos en la determinación de estrategias y políticas para el tratamiento de los desechos animales', *Investigación Económica*, 59(227), 111-141.

specon19

Galindo, L.M. and Montesillo, J.L. (1999) 'La demanda de agua potable en México: estimaciones preliminares', *Investigación Económica*, 59(227), 27-43.

specon20

González-Aréchiga Ramírez-Wiella, B., Venegas Martínez, F. and Díaz Tinoco, J. (2000) 'Riesgo de tasas de interés e inmunización por duración y convexidad con futuros: análisis local y de valor en riesgo', *Investigación económica*, 60(233), 77-112.

specon21

Gómez González, V., Drucker, A.G., Escalante Semerena, R., Rubio Leonel, O., Rosales Puc, E., Canul, V. and Magaña, A. (2000) 'Implicaciones económicas y ambientales de la normatividad ambiental aplicada al tratamiento de desechos en los rastros yucatecos', *Investigación Económica*, 60(231), 181-214.

specon22

Salcines, J.V. and Faíña, A. (2001) 'Un acercamiento mengeriano al concepto de meta-orden de necesidades: prioridad y jerarquía en Carl Menger', *Investigación Económica*, 37-67.

specon23

Delgado Wise, R. and Rodríguez Ramírez, H. (2002) 'El migrante frente a los desafíos del desarrollo local y regional en México', *Investigación Económica*, 62(240), 77-101.

specon24

Barceinas Paredes, F., Bara, R. and Luis, J. (2003) '¿Es rentable para el sector público subsidiar la educación en México?', *Investigación Económica*, 62(244), 141-163.

specon25

León León, M. J. (1999) 'Los modelos keynesianos de Hicks: evaluación y crítica', *Investigación Económica*, 59(229), 37-73.

specon26

Durán Romero, G. (2000) 'Desarrollo sostenible. Una revisión de los indicadores económicos de sustentabilidad', *Investigación Económica*, 60(231), 109-138.

specon27

Guzmán Gallangos, F. (2001) '¿Dónde trabajaban hombres y mujeres a principios de los noventa, y dónde seis años después? Segregación ocupacional por género, en México de 1991 a 1997', *Investigación Económica*, 61(236), 93-135.

specon28

Canudas González, R.D.C. (2001) 'La remergencia del capital humano en el marco de tres paradigmas', *Investigación Económica*, 61(235), 13-44.

specon29

Avella Alaminos, I. (2002) 'Dos momentos en la evolución de los tratados comerciales de México con Europa en el siglo XIX. El camino de la reciprocidad comercial', *Investigación Económica*, 62(240), 103-128.

specon30

Mendoza, J.E. (1999) 'Liberalización comercial y elasticidad del tipo de cambio real efectivo de las importaciones y exportaciones manufactureras mexicanas', *Investigación Económica*, 59(228), 97-142.

specon31

Rozo, C.A. (2000) 'La globalización: propuesta y paradojas. La experiencia de los países industriales', *Investigación Económica*, 60(234), 73-114.

specon32

García de León, A. (2001) 'La real compañía de Inglaterra y el tráfico negrero en el Veracruz del siglo XVIII, 1713-1748', *Investigación Económica*, 61(237), 153-182.

specon33

Mora Contreras, J. (2002) 'El poder de las grandes empresas petroleras: 1900-1950', *Investigación Económica*, 62(241), 153-183.

specon34

Michel Padilla, R. (2003) 'México: la economía de los sectores institucionales, distribución del ingreso, ahorro y crecimiento', *Investigación Económica*, 63(243), 155-211.

specon35

De Alba, E., Barraza, M., Artaloitia, B. and Balmes, T. (1999) 'Análisis costo-efectividad del efecto del proyecto piloto de nutrición, alimentación y salud', *El Trimestre Económico*, 66(261), 35-63.

specon36

Werner, A.M. and Bazdresch, S. (2000) 'La inflación moderada y la inestabilidad del proceso inflacionario', *El Trimestre Económico*, 67(268), 487-522.

specon37

Esquivel, G. and Larraín B.F. (2000) 'Determinantes de las crisis cambiarias', *El Trimestre Económico*, 67(266), 191-237.

specon38

Mizala, A. (2001) 'Factores socioeconómicos explicativos de los resultados escolares en la educación secundaria en Chile', *El Trimestre Económico*, 68(272), 515-549.

specon39

Lopez-Calva, L. and Rivas, L.A. (2002) 'Acumulación de capital trabajo infantil y educación obligatoria', *El Trimestre Económico*, 69(273), 3-36.

specon40

Barceñas Paredes, F. and Raymond-Bara, J.L. (2003) 'Hipótesis de señalización versus capital humano. El caso de México', *El Trimestre Económico*, 70(277), 167-194.

specon41

Acevedo, G.L. (1999) 'Tendencias del financiamiento de empresas en México', *El Trimestre Económico*, 66(261), 112-141.

specon42

González, L.M. (1999) 'Cambios en la estructura salarial de México en el periodo 1988-1993 y el aumento en el rendimiento de la educación superior', *El Trimestre Económico*, 66(261), 189-226.

specon43

Fernández A.G. (2001) 'Análisis empírico de la sostenibilidad externa: el caso de Ecuador', *El Trimestre Económico*, 68(269), 135-153.

specon44

López-Acevedo, G. (2002) 'Incentivos y desarrollo profesional de los profesores en las escuelas mexicanas', *El Trimestre Económico*, 69(275), 389-434.

specon45

Guerrero de Lizardi, C. (2003) 'Modelo de crecimiento económico restringido por la balanza de pagos: evidencia para México, 1940-2000', *El Trimestre Económico*, 70(278), 253-273.

specon46

González, L.M. (2003) 'Apertura comercial y cambio tecnológico: efectos en el mercado laboral mexicano', *El Trimestre Económico*, 70(279), 457-505.

specon47

Ferreiro Aparicio, J. (1999) 'Análisis teórico de la interrelación entre la entrada de inversión extranjera directa y la estructura productiva de las regiones anfitrionas. El caso de las regiones industriales en declive', *El Trimestre Económico*, 66(261), 91-111.

specon48

Cáceres, L.R. (2000) 'Distribución del ingreso e integración económica', *El Trimestre Económico*, 67(266), 301-319.

specon49

Rivas, L. (2001) 'Impuestos de tasa fija y composición del gasto público. Algunas consecuencias para el crecimiento a largo plazo', *El Trimestre Económico*, 68(271), 307-338.

specon50

Castillo, R.A. (2002) 'La protección de los derechos de propiedad y la elección de la estructura de la propiedad corporativa', *El Trimestre Económico*, 69(273), 37-63.